(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 August 2003 (21.08.2003)

PCT

(10) International Publication Number
WO 03/068943 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US03/04712

(22) International Filing Date: 13 February 2003 (13.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/357,002	13 February 2002 (13.02.2002)	US
60/362,439	6 March 2002 (06.03.2002)	US
60/366,041	19 March 2002 (19.03.2002)	US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **LEHR-MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95014 (US). **KABLE, Amy, E.** [US/US]; 2345 Polk Street #4, San Francisco, CA 94109 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **MARQUIS, Joseph, P.** [US/US]; 4428 Lazy Lane, San Jose, CA 95135 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square, #712, Union City, CA 94587 (US). **TRAN, Uyen, K.** [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US). **JIN, Pei** [US/US]; 320 Curtner Avenue #D, Palo Alto, CA 94306 (US). **TANG, Y. , Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **ZEBARJADIAN, Yeganeh** [IR/US]; 830 Junipero Serra Boulevard, San Francisco, CA 94127 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue #5D, San Francisco, CA 94122 (US). **HAFALIA, April, J, A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **COCKS, Benjamin, G.** [AU/US]; 1215 Van Dyke Drive, Sunnyvale, CA 94087 (US). **WARREN, Bridget, A.** [US/US]; 1810 S. El Camino Real #B103, Encinitas, CA 94024 (US). **EMERLING, Brooke, M.** [US/US]; 833 W. Buena Avenue #2108, Chicago, IL 60613 (US). **PEARSON, Cecilia , I.** [US/US]; 1019 Forest Court, Palo Alto, CA 94301 (US). **CHIEN, David** [US/US]; 444 Yerba Buena Avenue, Los Altos, CA 94022 (US). **PETERSON, David, P.** [US/US]; 970 Cherry Avenue, San Jose, CA 95126 (US). **FU, Glenn, K.** [US/US]; 4215 Roscommon Way, Dublin, CA 94568 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **JACKSON, Alan, A.** [US/US]; 1541 Elwood Drive, Los Gatos, CA 95032 (US). **JIANG, Xin** [US/US]; 14371 Elva Avenue, Saratoga, CA 95070 (US). **HAWKINS, Phillip, R.** [US/US]; 750 North Shorline Boulevard #115, Mountain View, CA 94043 (US). **LAL, Preeti, G.** [US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **KHARE, Reena** [US/US]; 12650 Orella Court, Saratoga, CA 95070 (US). **LEE, Sally** [US/US]; 3643 26th Street, San Francisco, CA 94110 (US). **LEE, Soo, Yeun** [KR/US]; 40 Westdale Avenue, Daly City, CA 94015 (US). **RICHARDSON, Thomas, W.** [US/US]; 616 Canyon Road #107, Redwood City, CA 94062 (US). **CHANG, Hsin-Ru** [US/US]; 326 Treasure Island Drive, Belmont, CA 94002 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/068943 A2

(54) Title: SECRETED PROTEINS

(57) Abstract: Various embodiments of the invention provide human secreted proteins (SECP) and polynucleotides which identify and encode SECP. Embodiments of the invention also provide expression vectors, host cells, antibodies, agonists, and antagonists. Other embodiments provide methods for diagnosing,treating, or preventing disorders associated with aberrant expression of SECP.

SECRETED PROTEINS

TECHNICAL FIELD

The invention relates to novel nucleic acids, secreted proteins encoded by these nucleic acids, and to the use of these nucleic acids and proteins in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders. The invention also relates to the assessment of the effects of exogenous compounds on the expression of nucleic acids and secreted proteins.

BACKGROUND OF THE INVENTION

Protein transport and secretion are essential for cellular function. Protein transport is mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Proteins that are retained in the plasma membrane contain one or more transmembrane domains, each comprised of about 20 hydrophobic amino acid residues. Secreted proteins are generally synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include proteins with important roles in cell-to-cell signaling. Such proteins include transmembrane receptors and cell surface markers, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, enzymes, neuropeptides, vasomediators, cell surface markers, and antigen recognition molecules. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

Cell surface markers include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs

directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A.N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, MPs may be heavily glycosylated and may contain an Arginine-Glycine-Aspartate (RGD) tripeptide motif which may play a role in adhesive interactions. MPs include extracellular proteins such as fibronectin, collagen, galectin, vitronectin and its proteolytic derivative somatomedin B; and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) Kidney Int. 51:1413-1417; Sjaastad, M.D. and W.J. Nelson (1997) BioEssays 19:47-55.)

Mucins are highly glycosylated glycoproteins that are the major structural component of the mucus gel. The physiological functions of mucins are cytoprotection, mechanical protection, maintenance of viscosity in secretions, and cellular recognition. MUC6 is a human gastric mucin that is also found in gall bladder, pancreas, seminal vesicles, and female reproductive tract (Toribara, N.W. et al. (1997) J. Biol. Chem. 272:16398-16403). The MUC6 gene has been mapped to human chromosome 11 (Toribara, N.W. et al. (1993) J. Biol. Chem. 268:5879-5885). Hemomucin is a novel *Drosophila* surface mucin that may be involved in the induction of antibacterial effector molecules (Theopold, U. et al. (1996) J. Biol. Chem. 217:12708-12715).

Tuftelins are one of four different enamel matrix proteins that have been identified so far.

The other three known enamel matrix proteins are the amelogenins, enamelin and ameloblastin. Assembly of the enamel extracellular matrix from these component proteins is believed to be critical in producing a matrix competent to undergo mineral replacement (Paine, C.T. et al. (1998) Connect Tissue Res. 38:257-267). Tuftelin mRNA has been found to be expressed in human ameloblastoma tumor, a non-mineralized odontogenic tumor (Deutsch, D. et al. (1998) Connect. Tissue Res. 39:177-184).

Olfactomedin-related proteins are extracellular matrix, secreted glycoproteins with conserved C-terminal motifs. They are expressed in a wide variety of tissues and in a broad range of species, from *Caenorhabditis elegans* to *Homo sapiens*. Olfactomedin-related proteins comprise a gene family with at least 5 family members in humans. One of the five, TIGR/myocilin protein, is expressed in the eye and is associated with the pathogenesis of glaucoma (Kulkarni, N.H. et al. (2000) Genet. Res. 76:41-50). Research by Yokoyama, M. et al. (1996; DNA Res. 3:311-320) found a 135-amino acid protein, termed AMY, having 96% sequence identity with rat neuronal olfactomedin-releated ER localized protein in a neuroblastoma cell line cDNA library, suggesting an essential role for AMY in nerve tissue. Neuron-specific olfactomedin-related glycoproteins isolated from rat brain cDNA libraries show strong sequence similarity with olfactomedin. This similarity is suggestive of a matrix-related function of these glycoproteins in neurons and neurosecretory cells (Danielson, P.E. et al. (1994) J. Neurosci. Res. 38:468-478).

Mac-2 binding protein is a 90-kD serum protein (90K), a secreted glycoprotein isolated from both the human breast carcinoma cell line SK-BR-3, and human breast milk. It specifically binds to a human macrophage-associated lectin, Mac-2. Structurally, the mature protein is 567 amino acids in length and is proceeded by an 18-amino acid leader. There are 16 cysteines and seven potential N-linked glycosylation sites. The first 106 amino acids represent a domain very similar to an ancient protein superfamily defined by a macrophage scavenger receptor cysteine-rich domain (Koths, K. et al. (1993) J. Biol. Chem. 268:14245-14249). 90K is elevated in the serum of subpopulations of AIDS patients and is expressed at varying levels in primary tumor samples and tumor cell lines. Ullrich, A. et al. (1994; J. Biol. Chem. 269:18401-18407) have demonstrated that 90K stimulates host defense systems and can induce interleukin-2 secretion. This immune stimulation is proposed to be a result of oncogenic transformation, viral infection or pathogenic invasion (Ullrich et al., *supra*).

Semaphorins are a large group of axonal guidance molecules consisting of at least 30 different members and are found in vertebrates, invertebrates, and even certain viruses. All semaphorins contain the sema domain which is approximately 500 amino acids in length. Neuropilin, a semaphorin receptor, has been shown to promote neurite outgrowth *in vitro*. The extracellular region of

neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of neuropilin have been suggested to have roles in protein-protein interactions and are thought to be involved in the binding of semaphorins through the sema and the C-terminal domains (reviewed in Raper, J.A. (2000) Curr. Opin. Neurobiol. 10:88-94). Plexins are neuronal cell surface molecules that mediate cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Plexins have been shown to be expressed in the receptors and neurons of particular sensory systems (Ohta, K. et al. (1995) Cell 14:1189-1199). There is evidence that suggests that some plexins function to control motor and CNS axon guidance in the developing nervous system. Plexins, which themselves contain complete semaphorin domains, may be both the ancestors of classical semaphorins and binding partners for semaphorins (Winberg, M.L. et al (1998) Cell 95:903-916).

Human pregnancy-specific beta 1-glycoprotein (PSG) is a family of closely related glycoproteins of molecular weights of 72 KDa, 64KDa, 62KDa, and 54KDa. Together with the carcinoembryonic antigen, they comprise a subfamily within the immunoglobulin superfamily (Plouzek, C.A. and J.Y. Chou, (1991) Endocrinology 129:950-958) Different subpopulations of PSG have been found to be produced by the trophoblasts of the human placenta, and the amnionic and chorionic membranes (Plouzek, C.A. et al. (1993) Placenta 14:277-285).

Torsion dystonia is an autosomal dominant movement disorder consisting of involuntary muscular contractions. The disorder has been linked to a 3-base pair mutation in the DYT-1 gene, which encodes torsin A (Ozelius, L.J. et al. (1997) Nat. Genet. 17:40-48). Torsin A bears significant homology to the Hsp100/Clp family of ATPase chaperones, which are conserved in humans, rats, mice, and *C. elegans*. Strong expression of DYT-1 in neuronal processes indicates a potential role for torsins in synaptic communication (Kustedjo, K. et al. (2000) J. Biol. Chem. 275:27933-27939 and Konakova M. et al. (2001) Arch. Neurol. 58:921-927).

Autocrine motility factor (AMF) is one of the motility cytokines regulating tumor cell migration; therefore identification of the signaling pathway coupled with it has critical importance. Autocrine motility factor receptor (AMFR) expression has been found to be associated with tumor progression in thymoma (Ohta Y. et al. (2000) Int. J. Oncol. 17:259-264). AMFR is a cell surface glycoprotein of molecular weight 78KDa.

Hormones are secreted molecules that travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical compositions and mechanisms of action, hormones can be grouped into two categories. One category includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these

molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines (epinephrine, norepinephrine) and histamine, and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotropic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864.)

Pro-opiomelanocortin (POMC) is the precursor polypeptide of corticotropin (ACTH), a hormone synthesized by the anterior pituitary gland, which functions in the stimulation of the adrenal cortex. POMC is also the precursor polypeptide of the hormone beta-lipotropin (beta-LPH). Each hormone includes smaller peptides with distinct biological activities: alpha-melanotropin (alpha-MSH) and corticotropin-like intermediate lobe peptide (CLIP) are formed from ACTH; gamma-lipotropin (gamma-LPH) and beta-endorphin are peptide components of beta-LPH; while beta-MSH is contained within gamma-LPH. Adrenal insufficiency due to ACTH deficiency, resulting from a genetic mutation in exons 2 and 3 of POMC results in an endocrine disorder characterized by early-onset obesity, adrenal insufficiency, and red hair pigmentation (Chretien, M. et al. (1979) Can. J. Biochem. 57:1111-1121; Krude, H. et al. (1998) Nat. Genet. 19:155-157; Online Mendelian Inheritance in Man (OMIM) 176830).

Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with membrane proteins for activity. Complex interactions among these factors and their receptors trigger intracellular signal transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth factors such as epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, and platelet-derived growth factor. The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs). Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin, which function as hormones to regulate cellular functions other than proliferation.

Growth and differentiation factors play critical roles in neoplastic transformation of cells *in vitro* and in tumor progression *in vivo*. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both *in vivo* and *in vitro*. Moreover, some growth factors and growth factor receptors are related both structurally and functionally to oncoproteins. In addition, growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)

The Slit protein, first identified in Drosophila, is critical in central nervous system midline formation and potentially in nervous tissue histogenesis and axonal pathfinding. Itoh et al. (1998; Brain Res. Mol. Brain Res. 62:175-186) have identified mammalian homologues of the slit gene (human Slit-1, Slit-2, Slit-3 and rat Slit-1). The encoded proteins are putative secreted proteins containing EGF-like motifs and leucine-rich repeats, both of which are conserved protein-protein interaction domains. Slit-1, -2, and -3 mRNAs are expressed in the brain, spinal cord, and thyroid, respectively (Itoh et al., *supra*). The Slit family of proteins are indicated to be functional ligands of glypican-1 in nervous tissue and it is suggested that their interactions may be critical in certain stages during central nervous system histogenesis (Liang, Y. et al. (1999) J. Biol. Chem. 274:17885-17892).

Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C.R. et al. (1985) Endocrine Physiology, Oxford University Press, New York, NY, pp. 57-62.)

NP/VMs are involved in numerous neurological and cardiovascular disorders. For example, neuropeptide Y is involved in hypertension, congestive heart failure, affective disorders, and appetite regulation. Somatostatin inhibits secretion of growth hormone and prolactin in the anterior pituitary, as well as inhibiting secretion in intestine, pancreatic acinar cells, and pancreatic beta-cells. A reduction in somatostatin levels has been reported in Alzheimer's disease and Parkinson's disease. Vasopressin acts in the kidney to increase water and sodium absorption, and in higher concentrations stimulates

contraction of vascular smooth muscle, platelet activation, and glycogen breakdown in the liver. Vasopressin and its analogues are used clinically to treat diabetes insipidus. Endothelin and angiotensin are involved in hypertension, and drugs, such as captopril, which reduce plasma levels of angiotensin, are used to reduce blood pressure (Watson, S. and S. Arkinstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 194; 252; 284; 55; 111).

Neuropeptides have also been shown to have roles in nociception (pain). Vasoactive intestinal peptide appears to play an important role in chronic neuropathic pain. Nociceptin, an endogenous ligand for for the opioid receptor-like 1 receptor, is thought to have a predominantly anti-nociceptive effect, and has been shown to have analgesic properties in different animal models of tonic or chronic pain (Dickinson, T. and S.M. Fleetwood-Walker (1998) Trends Pharmacol. Sci. 19:346-348).

Other proteins that contain signal peptides include secreted proteins with enzymatic activity. Such activity includes, for example, oxidoreductase/dehydrogenase activity, transferase activity, hydrolase activity, lyase activity, isomerase activity, or ligase activity. For example, matrix metalloproteinases are secreted hydrolytic enzymes that degrade the extracellular matrix and thus play an important role in tumor metastasis, tissue morphogenesis, and arthritis (Reponen, P. et al. (1995) Dev. Dyn. 202:388-396; Firestein, G.S. (1992) Curr. Opin. Rheumatol. 4:348-354; Ray, J.M. and W.G. Stetler-Stevenson (1994) Eur. Respir. J. 7:2062-2072; and Mignatti, P. and D.B. Rifkin (1993) Physiol. Rev. 73:161-195). Additional examples are the acetyl-CoA synthetases which activate acetate for use in lipid synthesis or energy generation (Luong, A. et al. (2000) J. Biol. Chem. 275:26458-26466). The result of acetyl-CoA synthetase activity is the formation of acetyl-CoA from acetate and CoA. Acetyl-CoA sythetases share a region of sequence similarity identified as the AMP-binding domain signature. Acetyl-CoA synthetase has been shown to be associated with hypertension (Toh, H. (1991) Protein Seq. Data Anal. 4:111-117; and Iwai, N. et al. (1994) Hypertension 23:375-380).

A number of isomerases catalyze steps in protein folding, phototransduction, and various anabolic and catabolic pathways. One class of isomerases is known as peptidyl-prolyl *cis-trans* isomerases (PPIases). PPIases catalyze the *cis* to *trans* isomerization of certain proline imidic bonds in proteins. Two families of PPIases are the FK506 binding proteins (FKBPs), and cyclophilins (CyPs). FKBPs bind the potent immunosuppressants FK506 and rapamycin, thereby inhibiting signaling pathways in T-cells. Specifically, the PPIase activity of FKBPs is inhibited by binding of FK506 or rapamycin. There are five members of the FKBP family which are named according to their calculated molecular masses (FKBP12, FKBP13, FKBP25, FKBP52, and FKBP65), and localized to different regions of the cell where they associate with different protein complexes (Coss, M. et al. (1995) J. Biol. Chem. 270:29336-29341; Schreiber, S.L. (1991) Science 251:283-287).

The peptidyl-prolyl isomerase activity of CyP may be part of the signaling pathway that leads to T-cell activation. CyP isomerase activity is associated with protein folding and protein trafficking, and may also be involved in assembly/disassembly of protein complexes and regulation of protein activity. For example, in *Drosophila*, the CyP NinaA is required for correct localization of rhodopsins, while a mammalian CyP (Cyp40) is part of the Hsp90/Hsc70 complex that binds steroid receptors. The mammalian CypA has been shown to bind the *gag* protein from human immunodeficiency virus 1 (HIV-1), an interaction that can be inhibited by cyclosporin. Since cyclosporin has potent anti-HIV-1 activity, CypA may play an essential function in HIV-1 replication. Finally, Cyp40 has been shown to bind and inactivate the transcription factor c-Myb, an effect that is reversed by cyclosporin. This effect implicates CyPs in the regulation of transcription, transformation, and differentiation (Bergsma, D.J. et al (1991) J. Biol. Chem. 266:23204-23214; Hunter, T. (1998) Cell 92:141-143; and Leverson, J.D. and S.A. Ness, (1998) Mol. Cell. 1:203-211).

Gamma-carboxyglutamic acid (Gla) proteins rich in proline (PRGPs) are members of a family of vitamin K-dependent single-pass integral membrane proteins. These proteins are characterized by an extracellular amino terminal domain of approximately 45 amino acids rich in Gla. The intracellular carboxyl terminal region contains one or two copies of the sequence PPXY, a motif present in a variety of proteins involved in such diverse cellular functions as signal transduction, cell cycle progression, and protein turnover (Kulman, J.D. et al. (2001) Proc. Natl. Acad. Sci. USA 98:1370-1375). The process of post-translational modification of glutamic residues to form Gla is Vitamin K-dependent carboxylation. Proteins which contain Gla include plasma proteins involved in blood coagulation. These proteins are prothrombin, proteins C, S, and Z, and coagulation factors VII, IX, and X. Osteocalcin (bone-Gla protein, BGP) and matrix Gla-protein (MGP) also contain Gla (Friedman, P.A. and C.T. Przysiecki (1987) Int. J. Biochem. 19:1-7; Vermeer, C. (1990) Biochem. J. 266:625-636).

Immunoglobulins

Antigen recognition molecules are key players in the sophisticated and complex immune systems which all vertebrates have developed to provide protection from viral, bacterial, fungal, and parasitic infections. A key feature of the immune system is its ability to distinguish foreign molecules, or antigens, from "self" molecules. This ability is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as lymphocytes, granulocytes, and monocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily, members of which contain one or more repeats of a conserved structural domain. This Ig domain is comprised of antiparallel β sheets joined by a disulfide bond in an arrangement called the Ig fold. The criteria for a

protein to be a member of the Ig superfamily is to have one or more Ig domains, which are regions of 70-110 amino acid residues in length homologous to either Ig variable-like (V) or Ig constant-like (C) domains. Members of the Ig superfamily include antibodies (Ab), T cell receptors (TCRs), class I and II major histocompatibility (MHC) proteins and immune cell-specific surface markers such as the "cluster of differentiation" or CD antigens, CD2, CD3, CD4, CD8, poly-Ig receptors, Fc receptors, neural cell-adhesion molecule (NCAM) and platelet-derived growth factor receptor (PDGFR).

Ig domains (V and C) are regions of conserved amino acid residues that give a polypeptide a globular tertiary structure called an immunoglobulin (or antibody) fold, which consists of two approximately parallel layers of β-sheets. Conserved cysteine residues form an intrachain disulfide-bonded loop, 55-75 amino acid residues in length, which connects the two layers of β-sheets. Each β-sheet has three or four anti-parallel β-strands of 5-10 amino acid residues. Hydrophobic and hydrophilic interactions of amino acid residues within the β-strands stabilize the Ig fold (hydrophobic on inward facing amino acid residues and hydrophilic on the amino acid residues in the outward facing portion of the strands). A V domain consists of a longer polypeptide than a C domain, with an additional pair of β-strands in the Ig fold.

A consistent feature of Ig superfamily genes is that each sequence of an Ig domain is encoded by a single exon. It is possible that the superfamily evolved from a gene coding for a single Ig domain involved in mediating cell-cell interactions. New members of the superfamily then arose by exon and gene duplications. Modern Ig superfamily proteins contain different numbers of V and/or C domains. Another evolutionary feature of this superfamily is the ability to undergo DNA rearrangements, a unique feature retained by the antigen receptor members of the family.

Many members of the Ig superfamily are integral plasma membrane proteins with extracellular Ig domains. The hydrophobic amino acid residues of their transmembrane domains and their cytoplasmic tails are very diverse, with little or no homology among Ig family members or to known signal-transducing structures. There are exceptions to this general superfamily description. For example, the cytoplasmic tail of PDGFR has tyrosine kinase activity. In addition Thy-1 is a glycoprotein found on thymocytes and T cells. This protein has no cytoplasmic tail, but is instead attached to the plasma membrane by a covalent glycophosphatidylinositol linkage.

Another common feature of many Ig superfamily proteins is the interactions between Ig domains which are essential for the function of these molecules. Interactions between Ig domains of a multimeric protein can be either homophilic or heterophilic (i.e., between the same or different Ig domains). Antibodies are multimeric proteins which have both homophilic and heterophilic interactions between Ig domains. Pairing of constant regions of heavy chains forms the Fc region of an antibody

and pairing of variable regions of light and heavy chains form the antigen binding site of an antibody. Heterophilic interactions also occur between Ig domains of different molecules. These interactions provide adhesion between cells for significant cell-cell interactions in the immune system and in the developing and mature nervous system. (Reviewed in Abbas, A.K. et al. (1991) Cellular and Molecular Immunology, W.B. Saunders Company, Philadelphia, PA, pp. 142-145.)

Antibodies

MHC proteins are cell surface markers that bind to and present foreign antigens to T cells. MHC molecules are classified as either class I or class II. Class I MHC molecules (MHC I) are expressed on the surface of almost all cells and are involved in the presentation of antigen to cytotoxic T cells. For example, a cell infected with virus will degrade intracellular viral proteins and express the protein fragments bound to MHC I molecules on the cell surface. The MHC I/antigen complex is recognized by cytotoxic T-cells which destroy the infected cell and the virus within. Class II MHC molecules are expressed primarily on specialized antigen-presenting cells of the immune system, such as B-cells and macrophages. These cells ingest foreign proteins from the extracellular fluid and express MHC II/antigen complex on the cell surface. This complex activates helper T-cells, which then secrete cytokines and other factors that stimulate the immune response. MHC molecules also play an important role in organ rejection following transplantation. Rejection occurs when the recipient's T-cells respond to foreign MHC molecules on the transplanted organ in the same way as to self MHC molecules bound to foreign antigen. (Reviewed in Alberts et al., *supra*, pp. 1229-1246.)

Antibodies are multimeric members of the Ig superfamily which are either expressed on the surface of B-cells or secreted by B-cells into the circulation. Antibodies bind and neutralize foreign antigens in the blood and other extracellular fluids. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the α, δ, ε, γ, and μ H-chain types. There are two types of L-chains, κ and λ, either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H- and L-

chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site. (Reviewed in Alberts et al. *supra*, pp. 1206-1213; 1216-1217.)

Both H-chains and L-chains contain the repeated Ig domains of members of the Ig superfamily. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region.

The immune system is capable of recognizing and responding to any foreign molecule that enters the body. Therefore, the immune system must be armed with a full repertoire of antibodies against all potential antigens. Such antibody diversity is generated by somatic rearrangement of gene segments encoding variable and constant regions. These gene segments are joined together by site-specific recombination which occurs between highly conserved DNA sequences that flank each gene segment. Because there are hundreds of different gene segments, millions of unique genes can be generated combinatorially. In addition, imprecise joining of these segments and an unusually high rate of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

Expression profiling

Microarrays are analytical tools used in bioanalysis. A microarray has a plurality of molecules spatially distributed over, and stably associated with, the surface of a solid support. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as gene sequencing, monitoring gene expression, gene mapping, bacterial identification, drug discovery, and combinatorial chemistry.

One area in particular in which microarrays find use is in gene expression analysis. Array technology can provide a simple way to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, condition, disease, or disorder.

Breast Cancer

There are more than 180,000 new cases of breast cancer diagnosed each year, and the mortality rate for breast cancer approaches 10% of all deaths in females between the ages of 45-54 (K. Gish (1999) AWIS Magazine 28:7-10). However the survival rate based on early diagnosis of localized breast cancer is extremely high (97%), compared with the advanced stage of the disease in which the tumor has spread beyond the breast (22%). Current procedures for clinical breast examination are lacking in sensitivity and specificity, and efforts are underway to develop comprehensive gene expression profiles for breast cancer that may be used in conjunction with conventional screening methods to improve diagnosis and prognosis of this disease (Perou CM et al. (2000) Nature 406:747-752).

Breast cancer is a genetic disease commonly caused by mutations in cellular disease. Mutations in two genes, BRCA1 and BRCA2, are known to greatly predispose a woman to breast cancer and may be passed on from parents to children (Gish, supra). However, this type of hereditary breast cancer accounts for only about 5% to 9% of breast cancers, while the vast majority of breast cancer is due to noninherited mutations that occur in breast epithelial cells.

A good deal is already known about the expression of specific genes associated with breast cancer. For example, the relationship between expression of epidermal growth factor (EGF) and its receptor, EGFR, to human mammary carcinoma has been particularly well studied. (See Khazaie et al., supra, and references cited therein for a review of this area.) Overexpression of EGFR, particularly coupled with down-regulation of the estrogen receptor, is a marker of poor prognosis in breast cancer patients. In addition, EGFR expression in breast tumor metastases is frequently elevated relative to the primary tumor, suggesting that EGFR is involved in tumor progression and metastasis. This is supported by accumulating evidence that EGF has effects on cell functions related to metastatic potential, such as cell motility, chemotaxis, secretion and differentiation. Changes in expression of other members of the erbB receptor family, of which EGFR is one, have also been implicated in breast cancer. The abundance of erbB receptors, such as HER-2/neu, HER-3, and HER-4, and their ligands in breast cancer points to their functional importance in the pathogenesis of the disease, and may therefore provide targets for therapy of the disease (Bacus, SS et al. (1994) Am J Clin Pathol 102:S13-S24). Other known markers of breast cancer include a human secreted frizzled protein mRNA that is downregulated in breast tumors; the matrix G1a protein which is overexpressed is human breast carcinoma cells; Drg1 or RTP, a gene whose expression is diminished in colon, breast, and prostate tumors; maspin, a tumor suppressor gene downregulated in invasive breast carcinomas; and CaN19, a member of the S100 protein family, all of which are down regulated in mammary

carcinoma cells relative to normal mammary epithelial cells (Zhou Z et al. (1998) Int J Cancer 78:95-99; Chen, L et al. (1990) Oncogene 5:1391-1395; Ulrix W et al (1999) FEBS Lett 455:23-26; Sager, R et al. (1996) Curr Top Microbiol Immunol 213:51-64; and Lee, SW et al. (1992) Proc Natl Acad Sci USA 89:2504-2508).

Cell lines derived from human mammary epithelial cells at various stages of breast cancer provide a useful model to study the process of malignant transformation and tumor progression as it has been shown that these cell lines retain many of the properties of their parental tumors for lengthy culture periods (Wistuba II et al. (1998) Clin Cancer Res 4:2931-2938). Such a model is particularly useful for comparing phenotypic and molecular characteristics of human mammary epithelial cells at various stages of malignant transformation.

Colon Cancer

Colorectal cancer is the second leading cause of cancer deaths in the United States. Colon cancer is associated with aging, since 90% of the total cases occur in individuals over the age of 55. A widely accepted hypothesis is that several contributing genetic mutations must accumulate over time in an individual who develops the disease. To understand the nature of genetic alterations in colorectal cancer, a number of studies have focused on the inherited syndromes. The first known inherited syndrome, Familial Adenomatous Polyposis (FAP), is caused by mutations in the Adenomatous Polyposis Coli gene (APC), resulting in truncated or inactive forms of the protein. This tumor suppressor gene has been mapped to chromosome 5q. The second known inherited syndrome is hereditary nonpolyposis colorectal cancer (HNPCC), which is caused by mutations in mismatch repair genes.

Although hereditary colon cancer syndromes occur in a small percentage of the population and most colorectal cancers are considered sporadic, knowledge from studies of the hereditary syndromes can be generally applied. For instance, somatic mutations in APC occur in at least 80% of indiscriminate colon tumors. APC mutations are thought to be the initiating event in the disease. Other mutations occur subsequently. Approximately 50% of colorectal cancers contain activating mutations in ras, while 85% contain inactivating mutations in p53. Changes in these genes lead to gene expression changes in colon cancer. Less is understood about downstream targets of these mutations and the role they may play in cancer development and progression.

Tangier Disease

Tangier disease (TD) is a genetic disorder characterized by near absence of circulating high density lipoprotein (HDL) and the accumulation of cholesterol esters in many tissues, including tonsils, lymph nodes, liver, spleen, thymus, and intestine. Low levels of HDL represent a clear predictor of

premature coronary artery disease and homozygous TD correlates with a four- to six-fold increase in cardiovascular disease compared to controls. HDL plays a cardio-protective role in reverse cholesterol transport, the flux of cholesterol from peripheral cells such as tissue macrophages through plasma lipoproteins to the liver. The HDL protein, apolipoprotein A-I, plays a major role in this process, interacting with the cell surface to remove excess cholesterol and phospholipids. This pathway is severely impaired in TD and the defect lies in a specific gene, the ABC1 transporter. This gene is a member of the family of ATP-binding cassette transporters, which utilize ATP hydrolysis to transport a variety of substrates across membranes.

Cell lines

Cell lines derived from human mammary epithelial cells at various stages of breast cancer provide a useful model to study the process of malignant transformation and tumor progression as it has been shown that these cell lines retain many of the properties of their parental tumors for lengthy culture periods (Wistuba, I.I. et al. (1998) Clin. Cancer Res. 4:2931-2938). Such a model is particularly useful for comparing phenotypic and molecular characteristics of human mammary epithelial cells at various stages of malignant transformation.

The human C3A cell line is a clonal derivative of HepG2/C3 (hepatoma cell line, isolated from a 15-year-old male with liver tumor), which was selected for strong contact inhibition of growth. The use of a clonal population enhances the reproducibility of the cells. C3A cells have many characteristics of primary human hepatocytes in culture: i) expression of insulin receptor and insulin-like growth factor II receptor; ii) secretion of a high ratio of serum albumin compared with α-fetoprotein iii) conversion of ammonia to urea and glutamine; iv) metabolism of aromatic amino acids; and v) proliferation in glucose-free and insulin-free medium. The C3A cell line is now well established as an in vitro model of the mature human liver (Mickelson et al. (1995) Hepatology 22:866-875; Nagendra et al. (1997) Am J Physiol 272:G408-G416).

Gemfibrozil is a fibric acid antilipemic agent that lowers serum triglycerides and produces favorable changes in lipoproteins. Gemfibrozil is effective in reducing the risk of coronary heart disease in men (Frick, M.H., et al. (1987) New Engl. J. Med; 317:1237-1245). The compound can inhibit peripheral lipolysis and decrease hepatic extraction of free fatty acids, which, decreases hepatic triglyceride production. Gemfibrozil also inhibits the synthesis and increases the clearance of apolipoprotein B, a carrier molecule for VLDL. Gemfibrozil has variable effects on LDL cholesterol. Although it causes moderate reductions in patients with type IIa hyperlipoproteinemia, changes in patients with either type IIb or type IV hyperlipoproteinemia are unpredictable. In general, the HMG-CoA reductase inhibitors are more effective than gemfibrozil in reducing LDL cholesterol. At the

molecular level gemfibozil may function as a peroxisome proliferator-activated receptor (PPAR) agonist. Gemfibrozil is rapidly and completely absorbed from the GI tract and undergoes enterohepatic recirculation. Gemfibrozil is metabolized by the liver and excreted by the kidneys, mainly as metabolites, one of which possesses pharmacologic activity. Gemfibozil causes peroxisome proliferation and hepatocarcinogenesis in rats, which is a cause for concern generally for fibric acid derivative drugs. In humans, fibric acid derivatives are known to increase the risk of gall bladder disease although gemfibrozil is better tolerated than other fibrates. The relative safety of gemfibrozil in humans compared to rodent species including rats may be attributed to differences in metabolism and clearance of the compound in different species (Dix, K.J., et al., (1999) Drug Metab. Distrib. 27 (1) 138-146; Thomas, B.F., et al., (1999) Drug Metab. Distrib. 27 (1) 147-157).

Lung cancer

Lung cancer is the leading cause of cancer death for men and the second leading cause of cancer death for women in the U.S. Lung cancers are divided into four histopathologically distinct groups. Three groups (squamous cell carcinoma, adenocarcinoma, and large cell carcinoma) are classified as non-small cell lung cancers (NSCLCs). The fourth group of cancers is referred to as small cell lung cancer (SCLC). Deletions on chromosome 3 are common in this disease and are thought to indicate the presence of a tumor suppressor gene in this region. Activating mutations in K-ras are commonly found in lung cancer and are the basis of one of the mouse models for the disease.

In nearly 80% of patients diagnosed with lung cancer, metastasis has already occurred. Most commonly lung cancers metastasize to pleura, brain, bone, pericardium, and liver. This adversely affects the overall five-year survival rate which is 37% for squamous carcinoma, 27% for adenocarcinoma and large cell carcinoma, and less than 1% for small cell carcinomas. Earlier diagnosis and an systematic approach to identification, staging, and treatment could positively affect patient outcome (DeVita et al. (1997) Cancer: Principles and Practice of Oncology, Lippincott-Raven, Philadelphia PA) and Fauci et al. (1998) Harrison's Principals of Internal Medicine, McGraw Hill, New York, NY).

Lung cancers progress through a series of morphologically distinct stages from hyperplasia to invasive carcinoma. Malignant lung cancers are divided into two groups comprising four histopathological classes. The nonsmall cell lung carcinoma (NSCLC) group includes squamous cell carcinomas, adenocarcinomas, and large cell carcinomas and accounts for about 70% of all lung cancer cases. Adenocarcinomas typically arise in the peripheral airways and often form mucin secreting glands. Squamous cell carcinomas typically arise in proximal airways. The histogenesis of squamous cell carcinomas may be related to chronic inflammation and injury to the bronchial

epithelium, leading to squamous metaplasia. The small cell lung carcinoma (SCLC) group accounts for about 20% of lung cancer cases. SCLCs typically arise in proximal airways and exhibit a number of paraneoplastic syndromes including inappropriate production of adrenocorticotropin and anti-diuretic hormone.

Lung cancer cells accumulate numerous genetic lesions, many of which are associated with cytologically visible chromosomal aberrations. The high frequency of chromosomal deletions associated with lung cancer may reflect the role of multiple tumor suppressor loci in the etiology of this disease. Several studies report deletions of regions of chromosome 11 in NSCLC (Bepler, G. and Garcia-Blanco, M.A. (1994) Proc. Natl. Acad. Sci. USA 91:5513-7; Iizuka, M., et al. (1995) Genes, Chromosomes & Cancer 13:40-46; Rasio, D. (1995) Cancer Research 55:3988-91). Deletions in other chromosome arms such as 3p, 9p and 17p are also common. Other frequently observed genetic lesions include overexpression of telomerase, activation of oncogenes such as K-ras and c-myc, and inactivation of tumor suppressor genes such as RB, p53 and p16 (Toomey, D. et al.(2001) Cancer 92:2648-57; Zajac-Kaye M. (2001) Lung Cancer 34:S43-6; Wright, G. et al. (2000) Current Opinion in Oncology 12:143-8; Kohno, T. and Yokota, J. (1999) Carcinogenesis 20:1403-10).

Ovarian Cancer

Ovarian cancer is the leading cause of death from a gynecologic cancer. The majority of ovarian cancers are derived from epithelial cells, and 70% of patients with epithelial ovarian cancers present with late-stage disease. As a result, the long-term survival rates for this disease is very low. Identification of early-stage markers for ovarian cancer would significantly increase the survival rate. The molecular events that lead to ovarian cancer are poorly understood. Some of the known aberrations include mutation of p53 and microsatellite instability. Since gene expression patterns are likely to vary when normal ovary is compared to ovarian tumors, examination of gene expression in these tissues to identify possible markers for ovarian cancer is particularly relevant to improving diagnosis, prognosis, and treatment of this disease.

Prostate Cancer

Prostate cancer is a common malignancy in men over the age of 50, and the incidence increases with age. In the US, there are approximately 132,000 newly diagnosed cases of prostate cancer and more than 33,000 deaths from the disorder each year.

Once cancer cells arise in the prostate, they are stimulated by testosterone to a more rapid growth. Thus, removal of the testes can indirectly reduce both rapid growth and metastasis of the cancer. Over 95 percent of prostatic cancers are adenocarcinomas which originate in the prostatic

acini. The remaining 5 percent are divided between squamous cell and transitional cell carcinomas, both of which arise in the prostatic ducts or other parts of the prostate gland.

As with most cancers, prostate cancer develops through a multistage progression ultimately resulting in an aggressive, metastatic phenotype. The initial step in tumor progression involves the hyperproliferation of normal luminal and/or basal epithelial cells that become hyperplastic and evolve into early-stage tumors. The early-stage tumors are localized in the prostate but eventually may metastasize, particularly to the bone, brain or lung. About 80% of these tumors remain responsive to androgen treatment, an important hormone controlling the growth of prostate epithelial cells. However, in its most advanced state, cancer growth becomes androgen-independent and there is currently no known treatment for this condition.

A primary diagnostic marker for prostate cancer is prostate specific antigen (PSA). PSA is a tissue-specific serine protease almost exclusively produced by prostatic epithelial cells. The quantity of PSA correlates with the number and volume of the prostatic epithelial cells, and consequently, the levels of PSA are an excellent indicator of abnormal prostate growth. Men with prostate cancer exhibit an early linear increase in PSA levels followed by an exponential increase prior to diagnosis. However, since PSA levels are also influenced by factors such as inflammation, androgen and other growth factors, some scientists maintain that changes in PSA levels are not useful in detecting individual cases of prostate cancer.

Current areas of cancer research provide additional prospects for markers as well as potential therapeutic targets for prostate cancer. Several growth factors have been shown to play a critical role in tumor development, growth, and progression. The growth factors Epidermal Growth Factor (EGF), Fibroblast Growth Factor (FGF), and Tumor Growth Factor alpha (TGFα) are important in the growth of normal as well as hyperproliferative prostate epithelial cells, particularly at early stages of tumor development and progression, and affect signaling pathways in these cells in various ways (Lin J et al. (1999) Cancer Res. 59:2891-2897; Putz T et al. (1999) Cancer Res 59:227-233). The TGF-β family of growth factors are generally expressed at increased levels in human cancers and the high expression levels in many cases correlates with advanced stages of malignancy and poor survival (Gold LI (1999) Crit Rev Oncog 10:303-360). Finally, there are human cell lines representing both the androgen-dependent stage of prostate cancer (LNCap) as well as the androgen-independent, hormone refractory stage of the disease (PC3 and DU-145) that have proved useful in studying gene expression patterns associated with the progression of prostate cancer, and the effects of cell treatments on these expressed genes (Chung TD (1999) Prostate 15:199-207).

Adipocyte Differentiation

The potential application of gene expression profiling is relevant to improving diagnosis, prognosis, and treatment of disease. For example, both the levels and sequences expressed in tissues from subjects with obesity or type II diabetes may be compared with the levels and sequences expressed in normal tissue.

The primary function of adipose tissue is the ability to store and release fat during periods of feeding and fasting. White adipose tissue is the major energy reserve in periods of fasting, and its reserve is mobilized during energy deprivation. Adipose tissue is one of the primary target tissues for insulin, and adipogenesis and insulin resistance are linked in type II diabetes, non-insulin dependent diabetes mellitus (NIDDM). Cytologically the conversion of a preadipocytes into mature adipocytes is characterized by deposition of fat droplets around the nuclei. The conversion process in vivo can be induced by thiazolidinediones (TZDs) and other PPARγ agonists (Adams et al. (1997) J. Clin. Invest. 100:3149-3153) which also lead to increased sensitivity to insulin and reduced plasma glucose and blood pressure.

Thiazolidinediones (TZDs) act as agonists for the peroxisome-proliferator-activated receptor gamma (PPARγ), a member of the nuclear hormone receptor superfamily. TZDs reduce hyperglycemia, hyperinsulinemia, and hypertension, in part by promoting glucose metabolism and inhibiting gluconeogenesis. Roles for PPARγ and its agonists have been demonstrated in a wide range of pathological conditions including diabetes, obesity, hypertension, atherosclerosis, polycystic ovarian syndrome, and cancers such as breast, prostate, liposarcoma, and colon cancer.

The mechanism by which TZDs and other PPARγ agonists enhance insulin sensitivity is not fully understood, but may involve the ability of PPARγ to promote adipogenesis. When ectopically expressed in cultured preadipocytes, PPARγ is a potent inducer of adipocyte differentiation. TZDs, in combination with insulin and other factors, can also enhance differentiation of human preadipocytes in culture (Adams et al. (1997) J. Clin. Invest. 100:3149-3153). The relative potency of different TZDs in promoting adipogenesis in vitro is proportional to both their insulin sensitizing effects in vivo, and their ability to bind and activate PPARγ in vitro. Interestingly, adipocytes derived from omental adipose depots are refractory to the effects of TZDs. It has therefore been suggested that the insulin sensitizing effects of TZDs may result from their ability to promote adipogenesis in subcutaneous adipose depots (Adams et al., ibid). Further, dominant negative mutations in the PPARγ gene have been identified in two non-obese subjects with severe insulin resistance, hypertension, and overt non-insulin dependent diabetes mellitus (NIDDM) (Barroso et al. (1998) Nature 402:880-883).

NIDDM is the most common form of diabetes mellitus, a chronic metabolic disease that affects 143 million people worldwide. NIDDM is characterized by abnormal glucose and lipid

metabolism that result from a combination of peripheral insulin resistance and defective insulin secretion. NIDDM has a complex, progressive etiology and a high degree of heritability. Numerous complications of diabetes including heart disease, stroke, renal failure, retinopathy, and peripheral neuropathy contribute to the high rate of morbidity and mortality.

At the molecular level, PPARγ functions as a ligand activated transcription factor. In the presence of ligand, PPARγ forms a heterodimer with the retinoid X receptor (RXR) which then activates transcription of target genes containing one or more copies of a PPARγ response element (PPRE). Many genes important in lipid storage and metabolism contain PPREs and have been identified as PPARγ targets, including PEPCK, aP2, LPL, ACS, and FAT-P (Auwerx, J. (1999) Diabetologia 42:1033-1049). Multiple ligands for PPARγ have been identified. These include a variety of fatty acid metabolites; synthetic drugs belonging to the TZD class, such as Pioglitazone and Rosiglitazone (BRL49653); and certain non-glitazone tyrosine analogs such as GI262570 and GW1929. The prostaglandin derivative 15-dPGJ2 is a potent endogenous ligand for PPARγ.

Expression of PPARγ is very high in adipose but barely detectable in skeletal muscle, the primary site for insulin stimulated glucose disposal in the body. PPARγ is also moderately expressed in large intestine, kidney, liver, vascular smooth muscle, hematopoietic cells, and macrophages. The high expression of PPARγ in adipose suggests that the insulin sensitizing effects of TZDs may result from alterations in the expression of one or more PPARγ regulated genes in adipose tissue. Identification of PPARγ target genes will contribute to better drug design and the development of novel therapeutic strategies for diabetes, obesity, and other conditions.

Systematic attempts to identify PPARγ target genes have been made in several rodent models of obesity and diabetes (Suzuki et al. (2000) Jpn. J. Pharmacol. 84:113-123; Way et al. (2001) Endocrinology 142:1269-1277). However, a serious drawback of the rodent gene expression studies is that significant differences exist between human and rodent models of adipogenesis, diabetes, and obesity (Taylor (1999) Cell 97:9-12; Gregoire et al. (1998) Physiol. Reviews 78:783-809). Therefore, an unbiased approach to identifying TZD regulated genes in primary cultures of human tissues is necessary to fully elucidate the molecular basis for diseases associated with PPARγ activity.

The majority of research in adipocyte biology to date has been done using transformed mouse preadipocyte cell lines. The culture condition, which stimulates mouse preadipocyte differentiation is different from that for inducing human primary preadipocyte differentiation. In addition, primary cells are diploid and may therefore reflect the <u>in vivo</u> context better than aneuploid cell lines. Understanding the gene expression profile during adipogenesis in human will lead to understanding the fundamental mechanism of adiposity regulation. Furthermore, through comparing the gene expression

profiles of adipogenesis between donor with normal weight and donor with obesity, identification of crucial genes, potential drug targets for obesity and type II diabetes, will be possible.

Dendritic Cells

Dendritic cells (DC) are antigen presenting cells (APC) that play a key role in the primary immune response because of their unique ability to present antigens to naive T cells. In addition, DC differentiate into separate subsets that sustain and regulate immune responses following initial contact with antigen. DC subsets include those that preferentially induce particular T helper 1 (Th1) or T helper 2 (Th2) responses and those that regulate B cell responses. Moreover, DC are increasingly being used to manipulate immune responses, either to downregulate an aberrant autoimmune response or to enhance vaccination or a tumor-specific response.

DC are functionally specialized in correlation with their particular differentiation state. CD34+ myeloid cells found in the bone marrow mature in response to as yet unclear signals into CD14+ CD11c+ monocytes. An innate or antigen non-specific response takes place initially when monocytes circulate to nonlymphoid tissues and respond to lipopolysaccharide (LPS), a bacterially-derived mitogen, and viruses. Such direct encounter with antigen causes secretion of pro-inflammatory cytokines that attract and regulate natural killer cells, macrophages, and eosinophils in the first line of defense against invading pathogens. Monocytes then mature into DC, which capture antigen highly efficiently through endocytosis and antigen receptor uptake. Antigen processing and presentation trigger activation and differentiation into mature DC that express MHC class II molecules on the cell surface and efficiently activate T cells, initiating antigen-specific T cell and B cell responses. In turn, T cells activate DC through CD40 ligand - CD40 interactions, which stimulate expression of the costimulatory molecules CD80 and CD86, the latter most potent in amplifying T cell responses. DC interaction via CD40 with T cells also stimulates the production of inflammatory cytokines such as TNF alpha and IL-1. Engagement of RANK, a member of the TNF receptor family by its ligand, TRANCE, which is expressed on activated T cells, enhances the survival of DC through inhibition of apoptosis, thereby enhancing T cell activation. The maturation and differentiation of monocytes into mature DC links the antigen non-specific innate immune response to the antigen-specific adaptive immune response.

The process by which monocytes differentiate into immature dendritic cells in vivo has not been fully elucidated. Incubation of monocytes with granulocyte-macrophage colony stimulating factor (GM-CSF) and interleukin (IL) -4 in vitro yields cells that exhibit functional and morphological characteristics equivalent to immature dendritic cells found in vivo. Moreover, incubation in vitro of immature dendritic cells with tumor necrosis factor alpha (TNF-α), CD40 ligand, LPS, or monocyte-

conditioned medium yields mature dendritic cells that are potent activators of naive T cells.

The ability to manipulate DC in vitro and their capacity to mount an effective immune response with small numbers of DC and little antigen has led to potential immunotherapies for diseases such as cancer, AIDS, and infectious diseases; and enhancing vaccine efficacy. Spontaneous remissions of particular cancers such as renal cell carcinomas and melanomas indicate that the immune system can respond to tumor antigens and eliminate tumors. However, tumors escape immune surveillance through a number of means including secretion of IL-10, macrophage colony stimulating factor, IL-6, and vascular endothelial growth factor, all of which inhibit DC activity and promote tolerance of tumor tissue. Delivery of tumor antigen-loaded DC to tumors can induce tumor-specific rejection in animal models. Similarly, pathogens can escape immune surveillance by altering antigen processing and presentation pathways or interfering with maturation of antigen presenting cells. Rather than providing resistance, DC can complicate infection by hosting latent viruses such as Kaposi's virus and cytomegalovirus, complicating infection. HIV-1 and measles virus particles are efficiently produced in DC. Vaccines against tumors or infectious pathogens could be improved by systemic or local administration of DC loaded with tumor antigens or attenuated viral particles or components, respectively.

The expression of killer-inhibitor regulatory molecules, chemokines, chemokine receptors, and proteinases have been identified in DC through sequencing of ESTs. Continuing this search may reveal new lymphocyte-binding and antigen-processing molecules, transmembrane and secretory products, and transcription factors that may help to explain the specialized features of DC and allow manipulation of the immune system.

Alzheimer's Disease

Characterization of region-specific gene expression in the human brain provides a context and background for molecular neurobiology on a variety of neurological disorders. For example, Alzheimer's disease (AD) is a progressive, neurodestructive process of the human neocortex, characterized by the deterioration of memory and higher cognitive function. A progressive and irreversible brain disorder, AD is characterized by three major pathogenic episodes involving (a) an aberrant processing and deposition of beta-amyloid precursor protein (betaAPP) to form neurotoxic beta-amyloid (betaA) peptides and an aggregated insoluble polymer of betaA that forms the senile plaque, (b) the establishment of intraneuronal neuritic tau pathology yielding widespread deposits of agyrophilic neurofibrillary tangles (NFT) and (c) the initiation and proliferation of a brain-specific inflammatory response. These three seemingly disperse attributes of AD etiopathogenesis are linked by the fact that proinflammatory microglia, reactive astrocytes and their associated cytokines and

chemokines are associated with the biology of the microtubule associated protein tau, betaA speciation and aggregation. Missense mutations in the presenilin genes PS1 and PS2, implicated in early onset familial AD, cause abnormal betaAPP processing with resultant overproduction of betaA42 and related neurotoxic peptides. Specific betaA fragments such as betaA42 can further potentiate proinflammatory mechanisms. Expression of the inducible oxidoreductase cyclooxygenase-2 and cytosolic phospholipase A2 (cPLA2) is strongly activated during cerebral ischemia and trauma, epilepsy and AD, indicating the induction of proinflammatory gene pathways as a response to brain injury. Neurotoxic metals such as aluminum and zinc, both implicated in AD etiopathogenesis, and arachidonic acid, a major metabolite of brain cPLA2 activity, each polymerize hyperphosphorylated tau to form NFT-like bundles. Studies have identified a reduced risk for AD in patients aged over 70 years who were previously treated with non-steroidal anti-inflammatory drugs for non-CNS afflictions that include arthritis. (For a review of the interrelationships between the mechanisms of PS1, PS2 and betaAPP gene expression, tau and betaA deposition and the induction, regulation and proliferation in AD of the neuroinflammatory response, see Lukiw, W.J, and Bazan, N.G. (2000) Neurochem. Res. 2000 25:1173-1184).

There is a need in the art for new compositions, including nucleic acids and proteins, for the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders.

SUMMARY OF THE INVENTION

Various embodiments of the invention provide purified polypeptides, secreted proteins, referred to collectively as 'SECP' and individually as 'SECP-1,' 'SECP-2,' 'SECP-3,' 'SECP-4,' 'SECP-5,' 'SECP-6,' 'SECP-7,' 'SECP-8,' 'SECP-9,' 'SECP-10,' 'SECP-11,' 'SECP-12,' 'SECP-13,' 'SECP-14,' 'SECP-15,' 'SECP-16,' 'SECP-17,' 'SECP-18,' 'SECP-19,' 'SECP-20,' 'SECP-21,' 'SECP-22,' 'SECP-23,' 'SECP-24,' 'SECP-25,' 'SECP-26,' 'SECP-27,' 'SECP-28,' 'SECP-29,' 'SECP-30,' 'SECP-31,' 'SECP-32,' 'SECP-33,' 'SECP-34,' 'SECP-35,' 'SECP-36,' 'SECP-37,' 'SECP-38,' 'SECP-39,' 'SECP-40,' 'SECP-41,' 'SECP-42,' 'SECP-43,' 'SECP-44,' 'SECP-45,' 'SECP-46,' 'SECP-47,' 'SECP-48,' 'SECP-49,' 'SECP-50,' and 'SECP-51' and methods for using these proteins and their encoding polynucleotides for the detection, diagnosis, and treatment of diseases and medical conditions. Embodiments also provide methods for utilizing the purified secreted proteins and/or their encoding polynucleotides for facilitating the drug discovery process, including determination of efficacy, dosage, toxicity, and pharmacology. Related embodiments provide methods for utilizing the purified secreted proteins and/or their encoding polynucleotides for investigating the

pathogenesis of diseases and medical conditions.

An embodiment provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. Another embodiment provides an isolated polypeptide comprising an amino acid sequence of SEQ ID NO:1-51.

Still another embodiment provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. In another embodiment, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-51. In an alternative embodiment, the polynucleotide is selected from the group consisting of SEQ ID NO:52-102.

Still another embodiment provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. Another embodiment provides a cell transformed with the recombinant polynucleotide. Yet another embodiment provides a transgenic organism comprising the recombinant polynucleotide.

Another embodiment provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least

90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Yet another embodiment provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51.

Still yet another embodiment provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In other embodiments, the polynucleotide can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence

complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex. In a related embodiment, the method can include detecting the amount of the hybridization complex. In still other embodiments, the probe can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Still yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof. In a related embodiment, the method can include detecting the amount of the amplified target polynucleotide or fragment thereof.

Another embodiment provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and a pharmaceutically acceptable excipient. In one embodiment, the composition can comprise an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. Other embodiments provide a method of treating a disease or condition associated with decreased or abnormal expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

Yet another embodiment provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a

naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. Another embodiment provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

Still yet another embodiment provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. Another embodiment provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment the composition.

Another embodiment provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the

polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

Yet another embodiment provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

Still yet another embodiment provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

Another embodiment provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs

under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide can comprise a fragment of a polynucleotide selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for full length polynucleotide and polypeptide embodiments of the invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog, and the PROTEOME database identification numbers and annotations of PROTEOME database homologs, for polypeptides of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide embodiments, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide embodiments, along with selected fragments of the polynucleotides.

Table 5 shows representative cDNA libraries for polynucleotide embodiments.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze polynucleotides and polypeptides, along with applicable descriptions, references, and threshold parameters.

Table 8 shows single nucleotide polymorphisms found in polynucleotide sequences of the invention, along with allele frequencies in different human populations.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleic acids, and methods are described, it is understood that embodiments of the invention are not limited to the particular machines, instruments, materials, and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the invention.

As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with various embodiments of the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"SECP" refers to the amino acid sequences of substantially purified SECP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of SECP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

An "allelic variant" is an alternative form of the gene encoding SECP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides.

Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding SECP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as SECP or a polypeptide with at least one functional characteristic of SECP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding SECP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide encoding SECP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent SECP. Deliberate amino acid substitutions may be made on the basis of one or more similarities in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of SECP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" can refer to an oligopeptide, a peptide, a polypeptide, or a protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid. Amplification may be carried out using polymerase chain reaction (PCR) technologies or other nucleic acid amplification technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of SECP. Antagonists may include proteins such as antibodies, anticalins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments

thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind SECP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH_2), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker (Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13).

The term "intramer" refers to an aptamer which is expressed *in vivo*. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl. Acad. Sci. USA 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on

substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a polynucleotide having a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic SECP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide" and a "composition comprising a given polypeptide" can refer to any composition containing the given polynucleotide or polypeptide. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotides encoding SECP or fragments of SECP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer

program for fragment assembly, such as the GELVIEW fragment assembly system (Accelrys, Burlington MA) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is

one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of SECP or a polynucleotide encoding SECP which can be identical in sequence to, but shorter in length than, the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from about 5 to about 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:52-102 can comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:52-102, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:52-102 can be employed in one or more embodiments of methods of the invention, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:52-102 from related polynucleotides. The precise length of a fragment of SEQ ID NO:52-102 and the region of SEQ ID NO:52-102 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-51 is encoded by a fragment of SEQ ID NO:52-102. A fragment of SEQ ID NO:1-51 can comprise a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-51. For example, a fragment of SEQ ID NO:1-51 can be used as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-51. The precise length of a fragment of SEQ ID NO:1-51 and the region of SEQ ID NO:1-51 to which the fragment corresponds can be determined based on the intended purpose for the fragment using one or more analytical methods described herein or otherwise known in the art.

A "full length" polynucleotide is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, alternatively, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of identical nucleotide matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using one or more computer algorithms or programs known in the art or described herein. For example, percent identity can be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989; CABIOS 5:151-153) and in Higgins, D.G. et al. (1992; CABIOS 8:189-191). For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms which can be used is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other

polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of identical residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide. The

phrases "percent similarity" and "% similarity," as applied to polypeptide sequences, refer to the percentage of residue matches, including identical residue matches and conservative substitutions, between at least two polypeptide sequences aligned using a standardized algorithm. In contrast, conservative substitutions are not included in the calculation of percent identity between polypeptide sequences.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid

sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press, Cold Spring Harbor NY, ch. 9).

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is

strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acids by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid present in solution and another nucleic acid immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or polynucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of SECP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of SECP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, antibody, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of SECP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of SECP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding

sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an SECP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of SECP.

"Probe" refers to nucleic acids encoding SECP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acids. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in, for example, Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press, Cold Spring Harbor NY), Ausubel, F.M. et al. (1999; Short Protocols in Molecular Biology, 4th ed., John Wiley & Sons, New York NY), and Innis, M. et al. (1990; PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a nucleic acid that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook and Russell (*supra*). The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a

vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA molecule, is composed of the same linear sequence of nucleotides as the reference DNA molecule with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing SECP, nucleic acids encoding SECP, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably at least about 75% free, and most preferably at least about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters,

chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. In another embodiment, the nucleic acid can be introduced by infection with a recombinant viral vector, such as a lentiviral vector (Lois, C. et al. (2002) Science 295:868-872). The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook and Russell (*supra*).

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at

least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotides that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity or sequence similarity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity or sequence similarity over a certain defined length of one of the polypeptides.

THE INVENTION

Various embodiments of the invention include new human secreted proteins (SECP), the polynucleotides encoding SECP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide embodiments of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide

sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown. Column 6 shows the Incyte ID numbers of physical, full length clones corresponding to the polypeptide and polynucleotide sequences of the invention. The full length clones encode polypeptides which have at least 95% sequence identity to the polypeptide sequences shown in column 3.

Table 2 shows sequences with homology to polypeptide embodiments of the invention as identified by BLAST analysis against the GenBank protein (genpept) database and the PROTEOME database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog and the PROTEOME database identification numbers (PROTEOME ID NO:) of the nearest PROTEOME database homologs. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank and PROTEOME database homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites and glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Accelrys, Burlington MA), as well as amino acid residues comprising signature sequences, domains, and motifs, including the locations of signal peptides (as indicated by "Signal Peptide" and/or "signal_cleavage"). Column 5 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are secreted proteins. For example, SEQ ID NO:24 is 100% identical, from residue M1 to residue R110, to human MYG1 (probable metal-dependent hydrolase) homolog (GenBank ID g10444289) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 3.7e-58, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from BLAST, analysis of the PRODOM database provides further corroborative evidence that SEQ ID NO:24 is a

secreted hydrolase.

As another example, SEQ ID NO:46 is 70% identical, from residue G45 to residue P1317 and is 39% identical, from residue W13 to residue E351, to rat MEGF6 (GenBank ID g3449294) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability scores are both 0.0, which indicate the probabilities of obtaining the observed polypeptide sequence alignments by chance. As determined by BLAST analysis using the PROTEOME database, SEQ ID NO:46 has homology to rat MEGF6 (multiple epidermal growth factor (EGF)-like domains 6) which contains 30 epidermal growth factor-like motifs, is a putative secreted protein, and is predicted to bind calcium (PROTEOME ID: 662841|Egfl3). SEQ ID NO:46 also has homology to a human protein which contains five laminin epidermal growth factor (EGF)-like and 15 epidermal growth factor (EGF)-like domains (PROTEOME ID: 716683|MEGF11). SEQ ID NO:46 also contains EGF-like domains and laminin-type EGF-like domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM and SMART databases of conserved protein families/domains. (See Table 3.) Data from BLIMPS, MOTIFS, and additional BLAST analyses against PRODOM and DOMO databases provide further corroborative evidence that SEQ ID NO:46 is a protein containing EGF-like domains.

SEQ ID NO:1-23, SEQ ID NO:25-45, and SEQ ID NO:47-51 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-51 are described in Table 7.

As shown in Table 4, the full length polynucleotide embodiments were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide embodiments, and of fragments of the polynucleotides which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:52-102 or that distinguish between SEQ ID NO:52-102 and related polynucleotides.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank

cDNAs or ESTs which contributed to the assembly of the full length polynucleotides. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL_*XXXXXX*_N_1_N_2_*YYYYY*_N_3_N_4 represents a "stitched" sequence in which *XXXXXX* is the identification number of the cluster of sequences to which the algorithm was applied, and *YYYYY* is the number of the prediction generated by the algorithm, and $N_{1,2,3...}$, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FL*XXXXXX*_g*AAAAA*_g*BBBBB*_1_*N* is a "stretched" sequence, with *XXXXXX* being the Incyte project identification number, g*AAAAA* being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, g*BBBBB* being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, g*BBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	**Type of analysis and/or examples of programs**
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).

INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotides which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotides. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

Table 8 shows single nucleotide polymorphisms (SNPs) found in polynucleotide sequences of the invention, along with allele frequencies in different human populations. Columns 1 and 2 show the polynucleotide sequence identification number (SEQ ID NO:) and the corresponding Incyte project identification number (PID) for polynucleotides of the invention. Column 3 shows the Incyte identification number for the EST in which the SNP was detected (EST ID), and column 4 shows the identification number for the SNP (SNP ID). Column 5 shows the position within the EST sequence at which the SNP is located (EST SNP), and column 6 shows the position of the SNP within the full-length polynucleotide sequence (CB1 SNP). Column 7 shows the allele found in the EST sequence. Columns 8 and 9 show the two alleles found at the SNP site. Column 10 shows the amino acid encoded by the codon including the SNP site, based upon the allele found in the EST. Columns 11-14 show the frequency of allele 1 in four different human populations. An entry of n/d (not detected) indicates that the frequency of allele 1 in the population was too low to be detected, while n/a (not available) indicates that the allele frequency was not determined for the population.

The invention also encompasses SECP variants. Various embodiments of SECP variants can have at least about 80%, at least about 90%, or at least about 95% amino acid sequence identity to the SECP amino acid sequence, and can contain at least one functional or structural characteristic of SECP.

Various embodiments also encompass polynucleotides which encode SECP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:52-102, which encodes SECP. The polynucleotide

sequences of SEQ ID NO:52-102, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses variants of a polynucleotide encoding SECP. In particular, such a variant polynucleotide will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a polynucleotide encoding SECP. A particular aspect of the invention encompasses a variant of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NO:52-102 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:52-102. Any one of the polynucleotide variants described above can encode a polypeptide which contains at least one functional or structural characteristic of SECP.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide encoding SECP. A splice variant may have portions which have significant sequence identity to a polynucleotide encoding SECP, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to a polynucleotide encoding SECP over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide encoding SECP. For example, a polynucleotide comprising a sequence of SEQ ID NO:52 and a polynucleotide comprising a sequence of SEQ ID NO:53 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:54, a polynucleotide comprising a sequence of SEQ ID NO:55, a polynucleotide comprising a sequence of SEQ ID NO:56, and a polynucleotide comprising a sequence of SEQ ID NO:83 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:63 and a polynucleotide comprising a sequence of SEQ ID NO:64 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:66 and a polynucleotide comprising a sequence of SEQ ID NO:67 are splice variants of each other; and a polynucleotide comprising a sequence of SEQ ID NO:74 and a polynucleotide comprising a sequence of SEQ ID NO:75 are splice variants of each other; and a polynucleotide comprising a sequence of SEQ ID NO:96 and a polynucleotide comprising a sequence of SEQ ID NO:97 are splice variants of each other. Any one of the splice

variants described above can encode a polypeptide which contains at least one functional or structural characteristic of SECP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding SECP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring SECP, and all such variations are to be considered as being specifically disclosed.

Although polynucleotides which encode SECP and its variants are generally capable of hybridizing to polynucleotides encoding naturally occurring SECP under appropriately selected conditions of stringency, it may be advantageous to produce polynucleotides encoding SECP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SECP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of polynucleotides which encode SECP and SECP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic polynucleotide may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a polynucleotide encoding SECP or any fragment thereof.

Embodiments of the invention can also include polynucleotides that are capable of hybridizing to the claimed polynucleotides, and, in particular, to those having the sequences shown in SEQ ID NO:52-102 and fragments thereof, under various conditions of stringency (Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511). Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied

Biosystems), thermostable T7 polymerase (Amersham Biosciences, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Invitrogen, Carlsbad CA). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Amersham Biosciences), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art (Ausubel et al., *supra*, ch. 7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853).

The nucleic acids encoding SECP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art (Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include

sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotides or fragments thereof which encode SECP may be cloned in recombinant DNA molecules that direct expression of SECP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other polynucleotides which encode substantially the same or a functionally equivalent polypeptides may be produced and used to express SECP.

The polynucleotides of the invention can be engineered using methods generally known in the art in order to alter SECP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of SECP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These

preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, polynucleotides encoding SECP may be synthesized, in whole or in part, using one or more chemical methods well known in the art (Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232). Alternatively, SECP itself or a fragment thereof may be synthesized using chemical methods known in the art. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques (Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J.Y. et al. (1995) Science 269:202-204). Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of SECP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography (Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421). The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing (Creighton, *supra*, pp. 28-53).

In order to express a biologically active SECP, the polynucleotides encoding SECP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotides encoding SECP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of polynucleotides encoding SECP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where a polynucleotide sequence encoding SECP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control

signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

Methods which are well known to those skilled in the art may be used to construct expression vectors containing polynucleotides encoding SECP and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination (Sambrook and Russell, *supra*, ch. 1-4, and 8; Ausubel et al., *supra*, ch. 1, 3, and 15).

A variety of expression vector/host systems may be utilized to contain and express polynucleotides encoding SECP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems (Sambrook and Russell, *supra*; Ausubel et al., *supra*; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355). Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of polynucleotides to the targeted organ, tissue, or cell population (Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5:350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Buller, R.M. et al. (1985) Nature 317:813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31:219-226; Verma, I.M. and N. Somia (1997) Nature 389:239-242). The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotides encoding SECP. For example, routine cloning, subcloning, and propagation of polynucleotides encoding SECP can be achieved using a multifunctional *E. coli*

vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Invitrogen). Ligation of polynucleotides encoding SECP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509). When large quantities of SECP are needed, e.g. for the production of antibodies, vectors which direct high level expression of SECP may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of SECP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign polynucleotide sequences into the host genome for stable propagation (Ausubel et al., *supra*; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184).

Plant systems may also be used for expression of SECP. Transcription of polynucleotides encoding SECP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection (The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196).

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, polynucleotides encoding SECP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses SECP in host cells (Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors

may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355).

For long term production of recombinant proteins in mammalian systems, stable expression of SECP in cell lines is preferred. For example, polynucleotides encoding SECP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively (Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823). Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14). Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051). Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β-glucuronidase and its substrate β-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding SECP is inserted within a marker gene sequence, transformed cells containing

polynucleotides encoding SECP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding SECP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the polynucleotide encoding SECP and that express SECP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of SECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SECP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art (Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding SECP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, polynucleotides encoding SECP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Biosciences, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with polynucleotides encoding SECP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a

transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode SECP may be designed to contain signal sequences which direct secretion of SECP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted polynucleotides or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant polynucleotides encoding SECP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric SECP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of SECP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SECP encoding sequence and the heterologous protein sequence, so that SECP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In another embodiment, synthesis of radiolabeled SECP may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple

transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

SECP, fragments of SECP, or variants of SECP may be used to screen for compounds that specifically bind to SECP. One or more test compounds may be screened for specific binding to SECP. In various embodiments, 1, 2, 3, 4, 5, 10, 20, 50, 100, or 200 test compounds can be screened for specific binding to SECP. Examples of test compounds can include antibodies, anticalins, oligonucleotides, proteins (e.g., ligands or receptors), or small molecules.

In related embodiments, variants of SECP can be used to screen for binding of test compounds, such as antibodies, to SECP, a variant of SECP, or a combination of SECP and/or one or more variants SECP. In an embodiment, a variant of SECP can be used to screen for compounds that bind to a variant of SECP, but not to SECP having the exact sequence of a sequence of SEQ ID NO:1-51. SECP variants used to perform such screening can have a range of about 50% to about 99% sequence identity to SECP, with various embodiments having 60%, 70%, 75%, 80%, 85%, 90%, and 95% sequence identity.

In an embodiment, a compound identified in a screen for specific binding to SECP can be closely related to the natural ligand of SECP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner (Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2):Chapter 5). In another embodiment, the compound thus identified can be a natural ligand of a receptor SECP (Howard, A.D. et al. (2001) Trends Pharmacol. Sci.22:132-140; Wise, A. et al. (2002) Drug Discovery Today 7:235-246).

In other embodiments, a compound identified in a screen for specific binding to SECP can be closely related to the natural receptor to which SECP binds, at least a fragment of the receptor, or a fragment of the receptor including all or a portion of the ligand binding site or binding pocket. For example, the compound may be a receptor for SECP which is capable of propagating a signal, or a decoy receptor for SECP which is not capable of propagating a signal (Ashkenazi, A. and V.M. Divit (1999) Curr. Opin. Cell Biol. 11:255-260; Mantovani, A. et al. (2001) Trends Immunol. 22:328-336). The compound can be rationally designed using known techniques. Examples of such techniques include those used to construct the compound etanercept (ENBREL; Amgen Inc., Thousand Oaks CA), which is efficacious for treating rheumatoid arthritis in humans. Etanercept is an engineered p75 tumor necrosis factor (TNF) receptor dimer linked to the Fc portion of human IgG_1 (Taylor, P.C. et al. (2001) Curr. Opin. Immunol. 13:611-616).

In one embodiment, two or more antibodies having similar or, alternatively, different specificities can be screened for specific binding to SECP, fragments of SECP, or variants of SECP. The binding specificity of the antibodies thus screened can thereby be selected to identify particular fragments or variants of SECP. In one embodiment, an antibody can be selected such that its binding specificity allows for preferential identification of specific fragments or variants of SECP. In another embodiment, an antibody can be selected such that its binding specificity allows for preferential diagnosis of a specific disease or condition having increased, decreased, or otherwise abnormal production of SECP.

In an embodiment, anticalins can be screened for specific binding to SECP, fragments of SECP, or variants of SECP. Anticalins are ligand-binding proteins that have been constructed based on a lipocalin scaffold (Weiss, G.A. and H.B. Lowman (2000) Chem. Biol. 7:R177-R184; Skerra, A. (2001) J. Biotechnol. 74:257-275). The protein architecture of lipocalins can include a beta-barrel having eight antiparallel beta-strands, which supports four loops at its open end. These loops form the natural ligand-binding site of the lipocalins, a site which can be re-engineered *in vitro* by amino acid substitutions to impart novel binding specificities. The amino acid substitutions can be made using methods known in the art or described herein, and can include conservative substitutions (e.g., substitutions that do not alter binding specificity) or substitutions that modestly, moderately, or significantly alter binding specificity.

In one embodiment, screening for compounds which specifically bind to, stimulate, or inhibit SECP involves producing appropriate cells which express SECP, either as a secreted protein or on the cell membrane. Preferred cells can include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing SECP or cell membrane fractions which contain SECP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either SECP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with SECP, either in solution or affixed to a solid support, and detecting the binding of SECP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

An assay can be used to assess the ability of a compound to bind to its natural ligand and/or to

inhibit the binding of its natural ligand to its natural receptors. Examples of such assays include radio-labeling assays such as those described in U.S. Patent No. 5,914,236 and U.S. Patent No. 6,372,724. In a related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a receptor) to improve or alter its ability to bind to its natural ligands (Matthews, D.J. and J.A. Wells. (1994) Chem. Biol. 1:25-30). In another related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a ligand) to improve or alter its ability to bind to its natural receptors (Cunningham, B.C. and J.A. Wells (1991) Proc. Natl. Acad. Sci. USA 88:3407-3411; Lowman, H.B. et al. (1991) J. Biol. Chem. 266:10982-10988).

SECP, fragments of SECP, or variants of SECP may be used to screen for compounds that modulate the activity of SECP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for SECP activity, wherein SECP is combined with at least one test compound, and the activity of SECP in the presence of a test compound is compared with the activity of SECP in the absence of the test compound. A change in the activity of SECP in the presence of the test compound is indicative of a compound that modulates the activity of SECP. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising SECP under conditions suitable for SECP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of SECP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding SECP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease (see, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337). For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and

the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding SECP may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding SECP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding SECP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress SECP, e.g., by secreting SECP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of SECP and secreted proteins. In addition, examples of tissues expressing SECP promonocyte line treated with PMA, peripheral blood mononuclear cells treated with interleukin 10, breast carcinoma and primary mammary epithelial cells, and normal and tumorous colon tissues, cancerous and normal breast, colon, lung, ovarian, prostate tissues, adipocytes, peripheral blood mononuclear cells, dendritic cells, monocytes, C3A hepatoma cell line, Raji B lymphoblast cells, and human umbilical vein endothelial cells. Further examples of tissues expressing SECP can be found in Table 6 and can also be found in Example XI. Therefore, SECP appears to play a role in cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders. In the treatment of disorders associated with increased SECP expression or activity, it is desirable to decrease the expression or activity of SECP. In the treatment of disorders associated with decreased SECP expression or activity, it is desirable to increase the expression or activity of SECP.

Therefore, in one embodiment, SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease

(MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural

abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss.

In another embodiment, a vector capable of expressing SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified SECP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of SECP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those listed above.

In a further embodiment, an antagonist of SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP. Examples of such

disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders described above. In one aspect, an antibody which specifically binds SECP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express SECP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP including, but not limited to, those described above.

In other embodiments, any protein, agonist, antagonist, antibody, complementary sequence, or vector embodiments may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of SECP may be produced using methods which are generally known in the art. In particular, purified SECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SECP. Antibodies to SECP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. In an embodiment, neutralizing antibodies (i.e., those which inhibit dimer formation) can be used therapeutically. Single chain antibodies (e.g., from camels or llamas) may be potent enzyme inhibitors and may have application in the design of peptide mimetics, and in the development of immuno-adsorbents and biosensors (Muyldermans, S. (2001) J. Biotechnol. 74:277-302).

For the production of antibodies, various hosts including goats, rabbits, rats, mice, camels, dromedaries, llamas, humans, and others may be immunized by injection with SECP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to SECP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are substantially identical to a portion of the amino acid sequence of the natural protein. Short stretches of SECP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to SECP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SECP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for SECP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 246:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired

specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between SECP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering SECP epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for SECP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of SECP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple SECP epitopes, represents the average affinity, or avidity, of the antibodies for SECP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular SECP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the SECP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of SECP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of SECP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available (Catty, *supra*; Coligan et al., *supra*).

In another embodiment of the invention, polynucleotides encoding SECP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding

SECP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding SECP (Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press, Totawa NJ).

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein (Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102:469-475; Scanlon, K.J. et al. (1995) 9:1288-1296). Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors (Miller, A.D. (1990) Blood 76:271; Ausubel et al., *supra*; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63:323-347). Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art (Rossi, J.J. (1995) Br. Med. Bull. 51:217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87:1308-1315; Morris, M.C. et al. (1997) Nucleic Acids Res. 25:2730-2736).

In another embodiment of the invention, polynucleotides encoding SECP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in SECP expression or regulation causes disease, the expression of SECP from an appropriate population of transduced cells may alleviate the clinical manifestations

caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SECP are treated by constructing mammalian expression vectors encoding SECP and introducing these vectors by mechanical means into SECP-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J.-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of SECP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). SECP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding SECP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to SECP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding SECP under the control of an independent promoter or the retrovirus long

terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In an embodiment, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SECP to cells which have one or more genetic abnormalities with respect to the expression of SECP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999; Annu. Rev. Nutr. 19:511-544) and Verma, I.M. and N. Somia (1997; Nature 18:389:239-242).

In another embodiment, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding SECP to target cells which have one or more genetic abnormalities with respect to the expression of SECP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing SECP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has

been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999; J. Virol. 73:519-532) and Xu, H. et al. (1994; Dev. Biol. 163:152-161). The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another embodiment, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding SECP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for SECP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of SECP-coding RNAs and the synthesis of high levels of SECP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of SECP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can

be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of RNA molecules encoding SECP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA molecules encoding SECP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine,

guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

In other embodiments of the invention, the expression of one or more selected polynucleotides of the present invention can be altered, inhibited, decreased, or silenced using RNA interference (RNAi) or post-transcriptional gene silencing (PTGS) methods known in the art. RNAi is a post-transcriptional mode of gene silencing in which double-stranded RNA (dsRNA) introduced into a targeted cell specifically suppresses the expression of the homologous gene (i.e., the gene bearing the sequence complementary to the dsRNA). This effectively knocks out or substantially reduces the expression of the targeted gene. PTGS can also be accomplished by use of DNA or DNA fragments as well. RNAi methods are described by Fire, A. et al. (1998; Nature 391:806-811) and Gura, T. (2000; Nature 404:804-808). PTGS can also be initiated by introduction of a complementary segment of DNA into the selected tissue using gene delivery and/or viral vector delivery methods described herein or known in the art.

RNAi can be induced in mammalian cells by the use of small interfering RNA also known as siRNA. SiRNA are shorter segments of dsRNA (typically about 21 to 23 nucleotides in length) that result *in vivo* from cleavage of introduced dsRNA by the action of an endogenous ribonuclease. SiRNA appear to be the mediators of the RNAi effect in mammals. The most effective siRNAs appear to be 21 nucleotide dsRNAs with 2 nucleotide 3' overhangs. The use of siRNA for inducing RNAi in mammalian cells is described by Elbashir, S.M. et al. (2001; Nature 411:494-498).

SiRNA can either be generated indirectly by introduction of dsRNA into the targeted cell, or directly by mammalian transfection methods and agents described herein or known in the art (such as liposome-mediated transfection, viral vector methods, or other polynucleotide delivery/introductory methods). Suitable SiRNAs can be selected by examining a transcript of the target polynucleotide (e.g., mRNA) for nucleotide sequences downstream from the AUG start codon and recording the occurrence of each nucleotide and the 3' adjacent 19 to 23 nucleotides as potential siRNA target sites, with sequences having a 21 nucleotide length being preferred. Regions to be avoided for target siRNA sites include the 5' and 3' untranslated regions (UTRs) and regions near the start codon (within 75 bases), as these may be richer in regulatory protein binding sites. UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP endonuclease complex. The selected target sites for siRNA can then be compared to the appropriate genome database (e.g., human, etc.) using BLAST or other sequence comparison algorithms known in the art. Target sequences with significant homology to other coding sequences can be eliminated from consideration. The selected SiRNAs can be produced by chemical synthesis methods known in the art or by *in vitro*

transcription using commercially available methods and kits such as the SILENCER siRNA construction kit (Ambion, Austin TX).

In alternative embodiments, long-term gene silencing and/or RNAi effects can be induced in selected tissue using expression vectors that continuously express siRNA. This can be accomplished using expression vectors that are engineered to express hairpin RNAs (shRNAs) using methods known in the art (see, e.g., Brummelkamp, T.R. et al. (2002) Science 296:550-553; and Paddison, P.J. et al. (2002) Genes Dev. 16:948-958). In these and related embodiments, shRNAs can be delivered to target cells using expression vectors known in the art. An example of a suitable expression vector for delivery of siRNA is the PSILENCER1.0-U6 (circular) plasmid (Ambion). Once delivered to the target tissue, shRNAs are processed *in vivo* into siRNA-like molecules capable of carrying out gene-specific silencing.

In various embodiments, the expression levels of genes targeted by RNAi or PTGS methods can be determined by assays for mRNA and/or protein analysis. Expression levels of the mRNA of a targeted gene, can be determined by northern analysis methods using, for example, the NORTHERNMAX-GLY kit (Ambion); by microarray methods; by PCR methods; by real time PCR methods; and by other RNA/polynucleotide assays known in the art or described herein. Expression levels of the protein encoded by the targeted gene can be determined by Western analysis using standard techniques known in the art.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding SECP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased SECP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding SECP may be therapeutically useful, and in the treatment of disorders associated with decreased SECP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding SECP may be therapeutically useful.

In various embodiments, one or more test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method

commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding SECP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding SECP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding SECP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art (Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466).

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of SECP, antibodies to SECP, and mimetics, agonists, antagonists, or inhibitors of SECP.

In various embodiments, the compositions described herein, such as pharmaceutical compositions, may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery allows administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising SECP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, SECP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for

administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example SECP or fragments thereof, antibodies of SECP, and agonists, antagonists or inhibitors of SECP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind SECP may be used for the diagnosis of disorders characterized by expression of SECP, or in assays to monitor patients being treated with SECP or agonists, antagonists, or inhibitors of SECP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for SECP include methods which utilize the antibody and a label to detect SECP in human body

fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring SECP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of SECP expression. Normal or standard values for SECP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to SECP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of SECP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, polynucleotides encoding SECP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotides, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of SECP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of SECP, and to monitor regulation of SECP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotides, including genomic sequences, encoding SECP or closely related molecules may be used to identify nucleic acid sequences which encode SECP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding SECP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the SECP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:52-102 or from genomic sequences including promoters, enhancers, and introns of the SECP gene.

Means for producing specific hybridization probes for polynucleotides encoding SECP include the cloning of polynucleotides encoding SECP or SECP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter

groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotides encoding SECP may be used for the diagnosis of disorders associated with expression of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus; emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular

replacement, and coronary artery bypass graft surgery; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss. Polynucleotides encoding SECP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular embodiment, polynucleotides encoding SECP may be used in assays that detect the presence of associated disorders, particularly those mentioned above. Polynucleotides complementary to sequences encoding SECP may be labeled by standard methods and added to a

fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of polynucleotides encoding SECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of SECP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding SECP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding SECP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding SECP, or a fragment of a polynucleotide complementary to the polynucleotide encoding SECP, and will be employed under optimized conditions for identification of a specific gene or

condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from polynucleotides encoding SECP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from polynucleotides encoding SECP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

SNPs may be used to study the genetic basis of human disease. For example, at least 16 common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase pathway. Analysis of the distribution of SNPs in different populations is useful for investigating genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations

and their migrations (Taylor, J.G. et al. (2001) Trends Mol. Med. 7:507-512; Kwok, P.-Y. and Z. Gu (1999) Mol. Med. Today 5:538-543; Nowotny, P. et al. (2001) Curr. Opin. Neurobiol. 11:637-641).

Methods which may also be used to quantify the expression of SECP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves (Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, SECP, fragments of SECP, or antibodies specific for SECP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time (Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484; hereby expressly incorporated by reference herein). Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present

invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*, as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity (see, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm). Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In an embodiment, the toxicity of a test compound can be assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another embodiment relates to the use of the polypeptides disclosed herein to analyze the

proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of interest. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for SECP to quantify the levels of SECP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which

alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art (Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/25116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and thoroughly described in Schena, M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

In another embodiment of the invention, nucleic acid sequences encoding SECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-

355; Price, C.M. (1993) Blood Rev. 7:127-134; Trask, B.J. (1991) Trends Genet. 7:149-154). Once mapped, the nucleic acid sequences may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP) (Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357).

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data (Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968). Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding SECP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation (Gatti, R.A. et al. (1988) Nature 336:577-580). The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, SECP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between SECP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest (Geysen, et al. (1984) PCT application WO84/03564). In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with SECP, or fragments thereof, and washed. Bound SECP is then detected by methods well known in the art. Purified SECP can also be coated

directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding SECP specifically compete with a test compound for binding SECP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with SECP.

In additional embodiments, the nucleotide sequences which encode SECP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/357,002, U.S. Ser. No. 60/362,439, and U.S. Ser. No. 60/366,041, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Invitrogen), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the

POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Invitrogen), using the recommended procedures or similar methods known in the art (Ausubel et al., *supra*, ch. 5). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Biosciences) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Invitrogen, Carlsbad CA), PCDNA2.1 plasmid (Invitrogen), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Invitrogen.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Biosciences or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Amersham Biosciences); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (Ausubel et al., *supra*, ch. 7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Candida albicans* (Incyte Genomics, Palo Alto CA); hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM (Haft, D.H. et al. (2001) Nucleic Acids Res. 29:41-43); and HMM-based protein domain databases such as SMART (Schultz, J. et al. (1998) Proc. Natl. Acad. Sci. USA 95:5857-5864; Letunic, I. et al. (2002) Nucleic Acids Res. 30:242-244). (HMM is a probabilistic approach which analyzes consensus primary structures of gene families; see, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred,

Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM; and HMM-based protein domain databases such as SMART. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (MiraiBio, Alameda CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:52-102. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative secreted proteins were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94; Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a

FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode secreted proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted proteins. Potential secreted proteins were also identified by homology to Incyte cDNA sequences that had been annotated as secreted proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

<u>"Stitched" Sequences</u>

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent

type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:52-102 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:52-102 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances

are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

Association of SECP polynucleotides with Lung Cancer

Heritable forms of lung carcinoma have not been reported and thus, identification of relevant disease-associated genes through conventional linkage analysis is not possible. However, several studies of sporadic nonsmall cell lung carcinoma (NSCLC) tumors have reported loss of heterozygosity (LOH) in regions of chromosome 11 suggesting the presence of one or more tumor suppressor genes (Bepler, G. and Garcia-Blanco, M.A. (1994) Proc. Natl. Acad. Sci. 91:5513-5517; Iizuka, M. (1995) Genes, Chromosomes & Cancer 13:40-46; Rasio, D. (1995) Cancer Research 55:3988-91). In a study of 79 patients with lung cancer, Iizuka and coworkers found that 11q14-11q24.2 was deleted in many of the lung tumors studied. Mapping of this region with additional markers showed that the region of chromosome 11q bounded by markers D11S939 and D11S938 was commonly deleted (Iizuka, et al., *supra*). In another study it was shown that human A549 NSCLC cells transformed with a human-derived YAC clone containing a region of chromosome 11q within the region bounded by D11S939 and D11S938, exhibited little or no increase in cell number (versus control cells whose number increased 5-10-fold in the same 5 day period). Further studies of these hybrid cells showed a decrease in tumorigenicity and an increase in latency following injection into athymic, nude mice, as compared with mice injected with control A549 cells. These studies suggest the presence of a tumor suppressor gene within this region of chromosome 11q and support the association of LOH in this region with NSCLC.

Restriction fragment length polymorphism (RFLP) markers shown to be near regions of DNA known as sequence-tagged sites (STS), have been mapped to NT_Contigs generated by the Human Genome Project using ePCR (Schuler, G.D. (1997) Genome Research 7: 541-550, and (1998) Trends Biotechnol. 16(11):456-459). Contigs containing regions of DNA with known disease-associated markers are therefore used to identify SECP sequences that map to disease-associated regions of the genome.

Polynucleotides encoding SECP were mapped to NT_Contigs. Contigs longer than 1Mb were broken into subcontigs of 1Mb length with overlapping sections of 100kb. A preliminary step used an algorithm, similar to MEGABLAST, to define the mRNA sequence /masked genomic DNA contig

pairings. The cDNA/genomic pairings identified by the first algorithm were confirmed, and the SECP polynucleotides mapped to DNA contigs, using SIM4 (Florea, L. et al. (1998) Genome Res. 8:967-74, version May 2000) which had been optimized for high throughput processing and strand assignment confidence). The SIM4 output of the mRNA sequence/genomic contig pairs was further processed to determine the correct location of the SECP polynucleotides on the genomic contig, as well as their strand identity.

SEQ ID NO:63 was mapped to NT_Contig NT_009151_021.8 from Genbank, version 128, covering a 5.5 Mb region of the genome that also contains lung cancer-associated genetic markers D11S939 and D11S938. The maximum distance between SEQ ID NO:63 and markers D11S939 and D11S938, therefore, is 5.5 Mb. Thus, SEQ ID NO:63 is in proximity with genetic markers shown to consistently associate with lung cancer. Therefore, in various embodiments, SEQ ID NO:63 can be used for one or more of the following: i) determination of LOH in persons with lung cancer in the lung cancer disease region at 11q12-24.2, ii) diagnostic assays for lung cancer, and iii) developing therapeutics and/or other treatments for lung cancer.

SEQ ID NO:64 was mapped to NT_Contig NT_009151_021.8 from Genbank, version 128, covering a 5.5 Mb region of the genome that also contains lung cancer-associated genetic markers D11S939 and D11S938. The maximum distance between SEQ ID NO:64 and markers D11S939 and D11S938, therefore, is 5.5 Mb. Thus, SEQ ID NO:64 is in proximity with genetic markers shown to consistently associate with lung cancer. Therefore, in various embodiments, SEQ ID NO:64 can be used for one or more of the following: i) determination of LOH in persons with lung cancer in the lung cancer disease region at 11q12-24.2, ii) diagnostic assays for lung cancer, and iii) developing therapeutics and/or other treatments for lung cancer.

Association of SECP polynucleotides with Alzheimer's Disease

SECP polynucleotides were mapped to NT_Contigs, available from NCBI, using the following procedures. Contigs longer than 1Mb were broken into subcontigs of 1Mb in length with overlapping sections of 100kb. A preliminary step used an algorithm, similar to MEGABLAST, to define the mRNA sequence/masked genomic DNA contig pairings. The cDNA/genomic pairings identified by the first algorithm were run through Sim4 (Florea, L. et al. (1998) Genome Res. 8:967-74, version May 2000) that had been optimized in house for high throughput and strand assignment confidence). The Sim4 output of the mRNA sequence/genomic contig pairs was further processed to determine the correct location of the SECP polynucleotides on the genomic contig, and also their strand identity.

Loci on chromosomes that map to regions associated with particular diseases can be used as

markers for these particular diseases. These markers then can be used to develop diagnostic and therapeutic tools for these diseases. For example, loci on chromosome 10 are associated with or linked to Alzheimer's disease (AD), a progressive neurodegenerative disease that represents the most common form of dementia (Ait-Ghezala, G. et al. (2002) Neurosci Lett. 325:87-90). AD can be inherited as an autosomal dominant trait. Further, genetic studies have focused on identification of genes that are potential targets for new treatments or improved diagnostics. The deposition and aggregation of β-amyloid in specific regions of the brain are key neuropathological hallmarks of AD. Insulin-degrading enzyme (IDE) can degrade β-amyloid Abraham, R. et al. (2001) Hum. Genet. 109:646-652). The IDE gene has been mapped near an AD-associated locus, 10q23-q25 (Espinosa R. 3rd et al. (1991) Cytogenet. Cell Genet. 57:184-186). Linkage analysis using IDE gene markers was performed on 1426 subjects from 435 families in which at least two family members were affected with AD.

A logarithm of the odds ratio for linkage (lod) score of over 3 indicates a probability of 1 in 1000 that a particular marker was found solely by chance in affected individuals. Significant linkage (lod score of 3.3) was reported between the polymorphic marker D10S583, located at 115.3 cM on chromosome 10, and AD with age of onset ≥50 years (Betram, L. et al. (2000) Science 290:2302-2303). D10S583 maps 36 kb upstream of the IDE gene. Further analysis of this region, however, failed to show association of SNPs (single nucleotide polymorphisms) within the IDE gene and flanking regions with late-onset AD (LOAD), in a study of 134 Caucasian LOAD cases and 111 matched controls from the United Kingdom (Abraham, R. et al, supra). Thus, although the activity of IDE may not influence the susceptibility to LOAD, there is substantial linkage in the chromosomal region containing the IDE gene, marker D10S583, and AD. The IDE gene and D10S583 both map to contig NT_008769, which contains a region of chromosome 10 that is 9.16 Mb in size.

SEQ ID NO:61 mapped to a region of contig NT_008804_002.8 from GenBank (version 128), localizing SEQ ID NO:61 to within 9.16 Mb of the Alzheimer's disease locus on chromosome 10q. Thus, SEQ ID NO:61 is in proximity with loci shown to consistently associate with Alzheimer's disease. Therefore, in various embodiments, SEQ ID NO:61 can be used for one or more of the following: i) linkage analysis of persons and/or families to the AD disease region at 10q, ii) diagnostic assays for AD, and iii) developing therapeutics and/or other treatments for AD.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs

from a particular cell type or tissue have been bound (Sambrook and Russell, *supra*, ch. 7; Ausubel et al., *supra*, ch. 4).

Analogous computer techniques applying BLAST were used to search for identical or related molecules in databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{5 \text{ x minimum } \{\text{length(Seq. 1), length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotides encoding SECP are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all

categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding SECP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of SECP Encoding Polynucleotides

Full length polynucleotides are produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Biosciences), ELONGASE enzyme (Invitrogen), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the

concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Biosciences). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Biosciences), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Biosciences) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Biosciences) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotides are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Identification of Single Nucleotide Polymorphisms in SECP Encoding Polynucleotides

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:52-102 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of

basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants. An automated procedure of advanced chromosome analysis analysed the original chromatogram files in the vicinity of the putative SNP. Clone error filters used statistically generated algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

Certain SNPs were selected for further characterization by mass spectrometry using the high throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezualan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed no allelic variance in this population were not further tested in the other three populations.

X. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:52-102 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Biosciences), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Biosciences). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon

membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

XI. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing; see, e.g., Baldeschweiler et al., *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena, M., ed. (1999) DNA Microarrays: A Practical Approach, Oxford University Press, London). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements (Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31).

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is

reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Biosciences). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte Genomics). Specific control poly(A)$^+$ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech, Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Biosciences).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma-Aldrich, St. Louis MO) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1 μl of the array element DNA, at an average concentration of 100 ng/μl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2%

SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and

adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte Genomics). Array elements that exhibit at least about a two-fold change in expression, a signal-to-background ratio of at least about 2.5, and an element spot size of at least about 40%, are considered to be differentially expressed.

Expression

For example, SEQ ID NO:62, SEQ ID NO:74, and SEQ ID NO:75 showed increased expression in Tangier disease-derived fibroblasts compared to normal fibroblasts. In addition, both types of cells were cultured in the presence of cholesterol and compared with the same cell type cultured in the absence of cholesterol. Human fibroblasts were obtained from skin explants from both normal subjects and two patients with homozygous Tangier disease. Cell lines were immortalized by transfection with human papillomavirus 16 genes E6 and E7 and a neomycin resistance selectable marker. TD-derived cells are deficient in an assay of apoA-I mediated tritiated cholesterol efflux. Therefore, SEQ ID NO:62, SEQ ID NO:74, and SEQ ID NO:75 are useful in diagnostic assays for Tangier disease.

As another example, SEQ ID NO:62 showed decreased expression in lung tissue affected by adenocarcinoma versus uninvolved lung tissue, as determined by microarray analysis. Moderately differentiated adenocarcinoma of the right lung was compared to grossly uninvolved lung tissue from a 60 year-old donor (Huntsman Cancer Institute). Therefore, SEQ ID NO:62 is useful in monitoring treatment of, and diagnostic assays for, lung adenocarcinoma and other cell proliferative disorders.

As another example, SEQ ID NO:65 showed increased expression in colon tissue affected by adenocarcinoma versus uninvolved colon tissue, as determined by microarray analysis. Gene

expression profiles were obtained by comparing normal colon tissue from a 60-year-old donor to colon adenocarcinoma tumor tissue from the same donor (Huntsman Cancer Institute, Salt Lake City, UT). Therefore, SEQ ID NO:65 is useful in monitoring treatment of, and diagnostic assays for, colon adenocarcinoma and other cell proliferative disorders.

As another example, SEQ ID NO:69 showed increased expression in breast carcinoma cells versus a nonmalignant mammary epithelial cell line. The gene expression profile of a nonmalignant mammary epithelial cell line was compared to the gene expression profiles of breast carcinoma lines at different stages of tumor progression. Cell lines compared included: a) BT-20, a breast carcinoma cell line derived *in vitro* from the cells emigrating out of thin slices of tumor mass isolated from a 74-year-old female, b) BT-474, a breast ductal carcinoma cell line that was isolated from a solid, invasive ductal carcinoma of the breast obtained from a 60-year-old woman, c) BT-483, a breast ductal carcinoma cell line that was isolated from a papillary invasive ductal tumor obtained from a 23-year-old normal, menstruating, parous female with a family history of breast cancer, d) Hs 578T, a breast ductal carcinoma cell line isolated from a 74-year-old female with breast carcinoma, e) MCF7, a nonmalignant breast adenocarcinoma cell line isolated from the pleural effusion of a 69-year-old female, f) MCF-10A, a breast mammary gland (luminal ductal characteristics) cell line isolated. from a 36-year-old woman with fibrocystic breast disease, g) MDA-MB-468, a breast adenocarcinoma cell line isolated from the pleural effusion of a 51-year-old female with metastatic adenocarcinoma of the breast, and h) HMEC, a primary breast epithelial cell line isolated from a normal donor. Therefore, SEQ ID NO:69 is useful in monitoring treatment of, and diagnostic assays for, breast carcinoma and other cell proliferative disorders.

As another example, SEQ ID NO:70 showed decreased expression in C3A cells treated with gemfibrozil versus untreated cells. Early confluent C3A cells were treated with various amounts of Gemfibrozil (120, 600, 800, and 1200 μg/ml) dissolved CMC, for 1, 3, and 6 hours. Parallel samples of C3A cells were treated with 1% CMC only, as a control. Therefore, SEQ ID NO:70 is useful in monitoring treatment of, and diagnostic assays for, coronary heart disease and other autoimmune/inflammatory disorders.

As another example, SEQ ID NO:80 showed differential expression in cancerous tissues versus normal tissues as determined by microarray analysis. In one experiment, the expression of SEQ ID NO:80, as determined by microarray analysis, was decreased by at least two fold in breast tumor tissues relative to normal breast tissues. The breast tumor tissues were harvested from a 43 year old female donor diagnosed with invasive lobular carcinoma. The tumor is well differentiated and

metastatic. The normal breast tissues were harvested from grossly uninvolved breast tissue of the same donor. Therefore, SEQ ID NO:80 is useful as a diagnostic marker or as a potential therapeutic target for breast cancer.

In the second experiment, the expression of SEQ ID NO:80, as determined by microarray analysis, was increased by at least two fold in sigmoid colon tissues relative to normal sigmoid colon tissues. The sigmoid colon tumor tissue which originated from a metastatic gastric sarcoma (stromal tumor) was harvested from a 48 year old female donor. The normal sigmoid colon tissue was harvested from grossly uninvolved sigmoid colon tissue of the same donor. Therefore, SEQ ID NO:80 is useful as a diagnostic marker or as a potential therapeutic target for colon cancer.

In the third experiment, SEQ ID NO:80 also showed decreased expression in tissue affected by lung adenocarcinoma versus normal lung tissue as determined by microarray analysis. A sample of right lung tissue that showed moderately differentiated adenocarcinoma was compared to grossly uninvolved lung tissue from a 60 year old donor (Huntsman Cancer Institute, Salt Lake City, UT). Therefore, SEQ ID NO:80 is useful in diagnostic assays for and monitoring treatment of lung cancer.

In the fourth experiment, the expression of SEQ ID NO:80 was decreased by at least two fold in ovarian adenocarcinoma tissues relative to grossly uninvolved normal ovarian tissue from a 79 year old donor. Therefore, SEQ ID NO:80 is useful as a diagnostic marker or as a potential therapeutic target for ovarian cancer.

The four experiments described above showed that SEQ ID NO:80 is useful as a diagnostic marker or as a potential therapeutic target for breast, colon, lung, and ovarian cancers.

In an alternative example, the expression of SEQ ID NO:80 were increased by at least two-fold in treated human adipocytes from obese donors when compared to non-treated adipocytes from the same donors. The obese human primary subcutaneous preadipocytes were isolated from adipose tissue of a 36-year-old healthy female with a body mass index (BMI) of 27.7 (overweight but otherwise healthy). The preadipocytes were cultured and induced to differentiate into adipocytes by culturing them in the differentiation medium containing the active components, PPAR-γ agonist and human insulin. Human preadipocytes were treated with human insulin and PPAR-γ agonist for three days and subsequently were switched to medium containing insulin for 5, 9, and 12 more days before the cells were collected for analysis. Differentiated adipocytes were compared to untreated preadipocytes maintained in culture in the absence of inducing agents. An overall differentiation rate of more than 60% was observed after 15 days in culture. The experiment show that SEQ ID NO:80 is useful for the diagnosis, prognosis, or treatment of diabetes mellitus and other disorders, such as

obesity, hypertension, and atherosclerosis.

In another example, SEQ ID NO:81 showed differential expression in inflammatory responses as determined by microarray analysis. The expression of SEQ ID NO:81 was increased by at least two fold in peripheral blood mononuclear cells (PBMCs, 12% B lymphocytes, 40% T lymphocytes, 20% NK cells, 25% monocytes, and 3% various cells that include dendritic and progenitor cells) treated with 10 ng/ml recombinant interleukin 5 (IL-5) for 2 hours relative to untreated cells. IL-5 is a T cell-derived factor that promotes the proliferation, differentiation, and activation of eosinophils. IL-5 exerts its activity on target cells by binding to specific cell surface receptors. The functional high-affinity receptor for human IL-5 is composed of a low-affinity IL-5 binding a-subunit and a non-binding common b-subunit that is shared with the high-affinity receptors for GM-CSF and IL-3. Therefore, this experiment showed that SEQ ID NO:81 is useful in diagnostic assays for inflammatory responses.

In addition, the expression of SEQ ID NO:81 is up-regulated in immature dendritic cells relative to monocytes by at least three-fold. The types of cDNAs that are up-regulated in concert during the transition from monocyte to dendritic cell reflect DC's newly acquired functions, such as antigen uptake. This experiment also showed that SEQ ID NO:81 is useful in diagnostic assays for inflammatory responses.

Further, as determined by microarray analysis, SEQ ID NO:81 showed differential expression in BT483 breast carcinoma cell line versus HMEC primary mammary epithelial cells. BT483 is a breast ductal carcinoma cell line isolated from a papillary invasive ductal tumor from a 23-year-old normal, menstruating, parous female. HMEC, a primary mammary epithelial cell line was derived from normal human mammary tissue (Clonetics, San Diego, CA). The microarray experiments showed that the expression of SEQ ID NO:81 was increased by at least two fold in BT483 breast ductal carcinoma line grown either in the presence or the absence of growth factors or nutrients relative to HMEC primary mammary epithelial cells grown in the absence of growth factors and nutrients. Therefore, SEQ ID NO:81 is useful as diagnostic markers or as potential therapeutic targets for breast cancer.

SEQ ID NO:81 showed differential expression in prostate carcinoma cell lines versus normal prostate epithelial cells as determined by microarray analysis. Three prostate carcinoma cell lines, DU 145, LNCaP, and PC-3 were included in the experiments. DU 145 was isolated from a metastatic site in the brain of a 69 year old male with widespread metastatic prostate carcinoma. DU 145 has no detectable sensitivity to hormones; forms colonies in semi-solid medium; is only weekly positive for

acid phosphatase; and cells are negative for prostate specific antigen (PSA). LNCaP is a prostate carcinoma cell line isolated from a lymph node biopsy of a 50 year old male with metastatic prostate carcinoma. LNCaP expresses PSA, produces prostate acid phosphatase, and expresses androgen receptors. PC-3, a prostate adenocarcinoma cell line, was isolated from a metastatic site in the bone of a 62 year old male with grade IV prostate adenocarcinoma. The normal epithelial cell line, PrEC, is a primary prostate epithelial cell line isolated from a normal donor. In one experiment, the expression of cDNAs from the prostate carcinoma cell lines were compared to that of the normal prostate epithelial cells grown under the same conditions (in the absence of growth factors and hormones). This experiment showed that the expression of SEQ ID NO:81 was increased by at least two fold in both DU145 and LNCaP prostate carcinoma lines relative to PrECs. In the other experiment, the expression of cDNAs from the prostate carcinoma cell lines grown in optimal conditions (in the presence of growth factors and hormones) were compared to that of the normal prostate epithelial cells grown under restrictive conditions (in the absence of growth factors and hormones). The experiment showed that the expression of SEQ ID NO:81 was increased by at least two fold in DU145, LNCaP, and PC-3 prostate carcinoma lines relative to PrECs. Therefore, SEQ ID NO:81 is useful as a diagnostic marker or as a potential therapeutic target for prostate cancers.

In yet another example, SEQ ID NO:82 showed differential expression, as determined by microarray analysis, in liver C3A cells treated with insulin and LY294002. The human C3A cell line is a clonal derivative of HepG2/C3 (hepatoma cell line, isolated from a 15-year-old male with liver tumor), which was selected for strong contact inhibition of growth. The use of a clonal population enhances the reproducibility of the cells. C3A cells have many characteristics of primary human hepatocytes in culture: i) expression of insulin receptor and insulin-like growth factor II receptor; ii) secretion of a high ratio of serum albumin compared with α-fetoprotein iii) conversion of ammonia to urea and glutamine; iv) metabolism of aromatic amino acids; and v) proliferation in glucose-free and insulin-free medium. The C3A cell line is now well established as an <u>in vitro</u> model of the mature human liver (Mickelson et al. (1995) Hepatology 22:866-875; Nagendra et al. (1997) Am J Physiol 272:G408-G416). LY294002 is a PI3 kinase specific inhibitor that promotes cell cycle arrest of C3A cells and skews these cells towards a more 'liver-like' state. This factor also appears to enhance the metabolic activity of the cells, especially with respect to proteins such as P450. In this experiment, C3A cells were starved of insulin for 3 days (-3 d to 0 h) and then further cultured in the presence of both insulin and the PI3-K inhibitor LY924002 for 0 hour, 24 hours, 2 hours and 3 days. The result showed that the expression of SEQ ID NO:82 was increased by at least a two-fold at in expression at

three out of the four time points (0 hour, 24 hours, and 2 hours). This experiment indicates that SEQ ID NO:82 is useful diagnostic assays for liver diseases and as a potential biological marker and therapeutic agent in the treatment of liver diseases and disorders.

In an alternative example, the expression of SEQ ID NO:85 was decreased by at least two fold in peripheral blood mononuclear cells (PBMCs, 12% B lymphocytes, 40% T lymphocytes, 20% NK cells, 25% monocytes, and 3% various cells that include dendritic and progenitor cells) treated with 10 ng/ml recombinant interleukin 4 (IL-4) for 2 hours relative to untreated cells. IL-4 is a pleiotropic cytokine produced by activated T cells, mast cells, and basophils. It was initially identified as a B cell differentiation factor (BCDF) and a B cell stimulatory factor (BSF1). IL-4 exhibits anti-tumor effects both in vivo and in vitro. Recently, IL-4 was identified as an important regulator for the CD4+ subset (Th1-like vs. Th2-like) development. The biological effects of IL-4 are mediated by the binding of IL-4 to specific cell surface receptors. The functional high-affinity receptor for IL-4 consists of a ligand-binding subunit (IL-4 R) and a second subunit (b chain) that can modulate the ligand binding affinity of the receptor complex. In certain cell types, the gamma chain of the IL-2 receptor complex is a functional b chain of the IL-4 receptor complex. This experiment showed that SEQ ID NO:85 is useful in diagnostic assays for inflammatory responses.

In another example, SEQ ID NO:87 showed differential expression in inflammatory responses as determined by microarray analysis. In one experiment, Raji B lymphoblast cells were stimulated *in vitro* with 0.1 mM soluble PMA (phorbol 12-myristate 13-acetate) and 1 mg/ml ionomycin for 0.5, 1, 2, 4, and 8 hours. Treated cells were compared to untreated Raji cells kept in culture in the absence of stimuli. Raji is a B lymphoblast cell line (Burkitt's lymphoma) that was isolated from the left jaw of an 11-year-old male. This cell line tests positive for the Epstein-Barr Nuclear Antigen (EBNA) but does not carry any detectable virus particles. Raji has been extensively used to study signaling in human B cells, identify factors produced by human B cells, and study the anti-lymphoma immune response. PMA is a broad activator of the protein kinase C-dependent pathways. Ionomycin is a calcium ionophore that permits the entry of calcium in the cell, hence increasing the cytosolic calcium concentration. The combination of PMA and ionomycin activates two of the major signaling pathways used by mammalian cells to interact with their environment. In T cells, the combination of PMA and ionomycin mimics the type of secondary signaling events elicited during optimal B cell activation. This experiment showed that expression of SEQ ID NO:87 was increased by at least two fold in three out of five time points (1, 2, and 4 hours).

In another experiment, the expression of SEQ ID NO:87 was increased by at least two fold in

peripheral blood mononuclear cells (PBMCs, 12% B lymphocytes, 40% T lymphocytes, 20% NK cells, 25% monocytes, and 3% various cells that include dendritic and progenitor cells) treated with 10 ng/ml tumor necrosis factor alpha (TNF-α) for 2 hours relative to untreated cells. TNF-α is produced by neutrophils, activated lymphocytes, macrophages, NK cells, LAK cells, astrocytes, endothelial cells, smooth muscle cells, and some transformed cells. TNF-α occurs as a secreted, soluble form and a membrane-anchored form, both of which are biologically active. TNF-α plays a critical role in mediation of the inflammatory response and in mediation of resistance to infections and tumor growth.

In the third experiment, the expression of SEQ ID NO:87 was increased by at least two fold in peripheral blood mononuclear cells (PBMCs, 12% B lymphocytes, 40% T lymphocytes, 20% NK cells, 25% monocytes, and 3% various cells that include dendritic and progenitor cells) treated with lipopolysaccharide (LPS) for 2 or 4 hours relative to untreated cells. LPS elicits a variety of inflammatory responses, and because it activates complement by the alternative (properdin) pathway, it is often part of the pathology of gram-negative bacterial infections. It is thought that LPS released into the bloodstream by lysing gram-negative bacteria is first bound by certain plasma proteins identified as LPS-binding proteins. The LPS-binding protein complex interacts with CD14 receptors on monocytes, macrophages, B cells, and other types of receptors on endothelial cells. Activation of human B cells with LPS results in mitogenesis as well as immunoglobulin synthesis. In monocytes and macrophages three types of events are triggered during their interaction with LPS: 1) Production of cytokines, including IL-1, IL-6, IL-8, TNF-a, and platelet-activating factor. These in turn stimulate production of prostaglandins and leukotrienes, powerful mediators of inflammation and septic shock that accompany endotoxin toxemia. 2) Activation of the complement cascade. 3) Activation of the coagulation cascade.

Therefore, the three experiments described above showed that SEQ ID NO:87 is useful in diagnostic assays for inflammatory responses.

As a further example, the expression of SEQ ID NO:101 was increased by at least two fold in human umbilical vein endothelial cells (HUVECs) treated with 100 nM phorbol 12-myristate 13-acetate (PMA) and 10 ng/mltumor necrosis factor-alpha (TNF-α) for 24 hours relative to untreated HUVECs. Human umbilical vein endothelial cells are primary cells derived from the endothelium of the human umbilical vein and have been used as an experimental model for investigating the role of the endothelium in human vascular biology. PMA is an agonist of protein kinase C (PKC) which is a calcium-activated, phospho-lipid-dependent serine- and threonine-specific kinase that upon activation phosphorylates a broad range of secondary targets. TNF-α is a pleiotropic cytokine that plays a

central role in mediating the inflammatory response through activation of multiple signal transduction pathways. TNF-α is produced by activated lymphocytes, macrophages, and other white blood cells and can activate endothelial cells. Monitoring the endothelial cells' response to TNF-α at the level of mRNA expression can provide information necessary for better understanding of both TNF-α signaling pathways and endothelial cell biology. TNF-α is also known to cause translocation of PKC from the cytosol to the membrane where it phosphorylates a variety of targets. Inhibition of PKC in the TNF-α activation experimental model will help clarify the PKC-dependent events during TNF-α signaling. The experiment indicated that SEQ ID NO:101 is useful in diagnosis, prognosis, or treatment of inflammatory disorders and endothelial-related disorders, such as those defected in vascular tone regulation, coagulation, thrombosis, and atherosclerosis.

XII. Complementary Polynucleotides

Sequences complementary to the SECP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring SECP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of SECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the SECP-encoding transcript.

XIII. Expression of SECP

Expression and purification of SECP is achieved using bacterial or virus-based expression systems. For expression of SECP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express SECP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of SECP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding SECP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter

drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus (Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945).

In most expression systems, SECP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Biosciences). Following purification, the GST moiety can be proteolytically cleaved from SECP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). Purified SECP obtained by these methods can be used directly in the assays shown in Examples XVII, XVIII, XIX, and XX, where applicable.

XIV. Functional Assays

SECP function is assessed by expressing the sequences encoding SECP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT plasmid (Invitrogen, Carlsbad CA) and PCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium

iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994; Flow Cytometry, Oxford, New York NY).

The influence of SECP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding SECP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding SECP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XV. Production of SECP Specific Antibodies

SECP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize animals (e.g., rabbits, mice, etc.) and to produce antibodies using standard protocols.

Alternatively, the SECP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art (Ausubel et al., *supra*, ch. 11).

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity (Ausubel et al., *supra*). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-SECP activity by, for example, binding the peptide or SECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XVI. Purification of Naturally Occurring SECP Using Specific Antibodies

Naturally occurring or recombinant SECP is substantially purified by immunoaffinity

chromatography using antibodies specific for SECP. An immunoaffinity column is constructed by covalently coupling anti-SECP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Biosciences). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing SECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SECP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SECP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and SECP is collected.

XVII. Identification of Molecules Which Interact with SECP

SECP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent (Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled SECP, washed, and any wells with labeled SECP complex are assayed. Data obtained using different concentrations of SECP are used to calculate values for the number, affinity, and association of SECP with the candidate molecules.

Alternatively, molecules interacting with SECP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989; Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

SECP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVIII. Demonstration of SECP Activity

An assay for growth stimulating or inhibiting activity of SECP measures the amount of DNA synthesis in Swiss mouse 3T3 cells (McKay, I. and I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY). In this assay, varying amounts of SECP are added to quiescent 3T3 cultured cells in the presence of [^{3}H]thymidine, a radioactive DNA precursor. SECP for this assay can be obtained by recombinant means or from biochemical preparations. Incorporation of [^{3}H]thymidine into acid-precipitable DNA is measured over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold SECP concentration range is

indicative of growth modulating activity. One unit of activity per milliliter is defined as the concentration of SECP producing a 50% response level, where 100% represents maximal incorporation of [^{3}H]thymidine into acid-precipitable DNA .

Alternatively, an assay for SECP activity measures the stimulation or inhibition of neurotransmission in cultured cells. Cultured CHO fibroblasts are exposed to SECP. Following endocytic uptake of SECP, the cells are washed with fresh culture medium, and a whole cell voltage-clamped *Xenopus* myocyte is manipulated into contact with one of the fibroblasts in SECP-free medium. Membrane currents are recorded from the myocyte. Increased or decreased current relative to control values are indicative of neuromodulatory effects of SECP (Morimoto, T. et al. (1995) Neuron 15:689-696).

Alternatively, an assay for SECP activity measures the amount of SECP in secretory, membrane-bound organelles. Transfected cells as described above are harvested and lysed. The lysate is fractionated using methods known to those of skill in the art, for example, sucrose gradient ultracentrifugation. Such methods allow the isolation of subcellular components such as the Golgi apparatus, ER, small membrane-bound vesicles, and other secretory organelles. Immunoprecipitations from fractionated and total cell lysates are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The concentration of SECP in secretory organelles relative to SECP in total cell lysate is proportional to the amount of SECP in transit through the secretory pathway.

Alternatively, AMP binding activity is measured by combining SECP with ^{32}P-labeled AMP. The reaction is incubated at 37°C and terminated by addition of trichloroacetic acid. The acid extract is neutralized and subjected to gel electrophoresis to remove unbound label. The radioactivity retained in the gel is proportional to SECP activity.

XIX. Demonstration of Immunoglobulin Activity

An assay for SECP activity measures the ability of SECP to recognize and precipitate antigens from serum. This activity can be measured by the quantitative precipitin reaction. (Golub, E.S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pp. 113-115.) SECP is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled SECP. SECP-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable SECP-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable SECP-antigen complex is plotted against the serum concentration. For various serum concentrations, a characteristic

precipitin curve is obtained, in which the amount of precipitable SECP-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable SECP-antigen complex is a measure of SECP activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

Alternatively, an assay for SECP activity measures the expression of SECP on the cell surface. cDNA encoding SECP is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of SECP expressed on the cell surface.

Alternatively, an assay for SECP activity measures the amount of cell aggregation induced by overexpression of SECP. In this assay, cultured cells such as NIH3T3 are transfected with cDNA encoding SECP contained within a suitable mammalian expression vector under control of a strong promoter. Cotransfection with cDNA encoding a fluorescent marker protein, such as Green Fluorescent Protein (CLONTECH), is useful for identifying stable transfectants. The amount of cell agglutination, or clumping, associated with transfected cells is compared with that associated with untransfected cells. The amount of cell agglutination is a direct measure of SECP activity.

XX. SECP Secretion Assay

A high throughput assay may be used to identify polypeptides that are secreted in eukaryotic cells. In an example of such an assay, polypeptide expression libraries are constructed by fusing 5'-biased cDNAs to the 5'-end of a leaderless β-lactamase gene. β-lactamase is a convenient genetic reporter as it provides a high signal-to-noise ratio against low endogenous background activity and retains activity upon fusion to other proteins. A dual promoter system allows the expression of β-lactamase fusion polypeptides in bacteria or eukaryotic cells, using the *lac* or CMV promoter, respectively.

Libraries are first transformed into bacteria, *e.g.*, *E. coli*, to identify library members that encode fusion polypeptides capable of being secreted in a prokaryotic system. Mammalian signal sequences direct the translocation of β-lactamase fusion polypeptides into the periplasm of bacteria where they confer antibiotic resistance to carbenicillin. Carbenicillin-selected bacteria are isolated on solid media, individual clones are grown in liquid media, and the resulting cultures are used to isolate library member plasmid DNA.

Mammalian cells, *e.g.*, 293 cells, are seeded into 96-well tissue culture plates at a density of about 40,000 cells/well in 100 μl phenol red-free DME supplemented with 10% fetal bovine serum (FBS) (Life Technologies, Rockville, MD). The following day, purified plasmid DNAs isolated from carbenicillin-resistant bacteria are diluted with 15 μl OPTI-MEM I medium (Life Technologies) to a volume of 25 μl for each well of cells to be transfected. In separate plates, 1 μl LF2000 Reagent (Life Technologies) is diluted into 25 μl/well OPTI-MEM I. The 25 μl diluted LF2000 Reagent is then combined with the 25 μl diluted DNA, mixed briefly, and incubated for 20 minutes at room temperature. The resulting DNA-LF2000 reagent complexes are then added directly to each well of 293 cells. Cells are also transfected with appropriate control plasmids expressing either wild-type β-lactamase, leaderless β-lactamase, or, for example, CD4-fused leaderless β-lactamase. 24 hrs following transfection, about 90 μl of cell culture media are assayed at 37°C with 100 μM Nitrocefin (Calbiochem, San Diego, CA) and 0.5 mM oleic acid (Sigma Corp. St. Louis, MO) in 10 mM phosphate buffer (pH 7.0). Nitrocefin is a substrate for β-lactamase that undergoes a noticeable color change from yellow to red upon hydrolysis. β-lactamase activity is monitored over 20 min in a microtiter plate reader at 486 nm. Increased color absorption at 486 nm corresponds to secretion of a β-lactamase fusion polypeptide in the transfected cell media, resulting from the presence of a eukaryotic signal sequence in the fusion polypeptide. Polynucleotide sequence analysis of the corresponding library member plasmid DNA is then used to identify the signal sequence-encoding cDNA. (Described in U.S. Patent application 09/803,317, filed March 9, 2001.)

For example, SEQ ID NO:28 was shown to be a secreted protein using this assay.

Various modifications and variations of the described compositions, methods, and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. It will be appreciated that the invention provides novel and useful proteins, and their encoding polynucleotides, which can be used in the drug discovery process, as well as methods for using these compositions for the detection, diagnosis, and treatment of diseases and conditions. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Nor should the description of such embodiments be considered exhaustive or limit the invention to the precise forms disclosed. Furthermore, elements from one embodiment can be readily recombined with elements from one or more other embodiments. Such combinations can form a number of embodiments within the scope of the invention. It is intended that the scope of the invention be

defined by the following claims and their equivalents.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
7506904	1	7506904CD1	52	7506904CB1	3216178CA2, 5291455CA2, 90072791CA2, 90109312CA2, 90109351CA2, 90109375CA2, 90109403CA2, 90109435CA2, 90109459CA2, 90109467CA2, 90109511CA2, 90109550CA2, 90109559CA2, 90109603CA2, 90109627CA2, 90109635CA2, 90109643CA2, 90109651CA2, 90109659CA2, 90109673CA2, 90109675CA2, 90109682CA2, 90109691CA2
7506909	2	7506909CD1	53	7506909CB1	90109350CA2, 90109404CA2, 90109412CA2, 90109426CA2, 90109503CA2, 90109567CA2, 90109573CA2, 90109618CA2, 90109658CA2, 90109683CA2, 90109690CA2
7507096	3	7507096CD1	54	7507096CB1	90115488CA2, 90115532CA2
7507098	4	7507098CD1	55	7507098CB1	
7507099	5	7507099CD1	56	7507099CB1	
7501399	6	7501399CD1	57	7501399CB1	
7504768	7	7504768CD1	58	7504768CB1	4111686CA2
7500757	8	7500757CD1	59	7500757CB1	90158490CA2
1730616	9	1730616CD1	60	1730616CB1	
190404	10	190404CD1	61	190404CB1	
7500679	11	7500679CD1	62	7500679CB1	
7500687	12	7500687CD1	63	7500687CB1	
7500688	13	7500688CD1	64	7500688CB1	90207830CA2
7500697	14	7500697CD1	65	7500697CB1	4330768CA2, 5072466CA2
7500709	15	7500709CD1	66	7500709CB1	
7500711	16	7500711CD1	67	7500711CB1	90213704CA2
7500723	17	7500723CD1	68	7500723CB1	7978329CA2, 95074077CA2
7500764	18	7500764CD1	69	7500764CB1	7577131CA2, 90197034CA2, 90197102CA2
7500772	19	7500772CD1	70	7500772CB1	
7501350	20	7501350CD1	71	7501350CB1	90197157CA2, 90203017CA2
7506396	21	7506396CD1	72	7506396CB1	
7505917	22	7505917CD1	73	7505917CB1	

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
7500701	23	7500701CD1	74	7500701CB1	
7500702	24	7500702CD1	75	7500702CB1	90210690CA2
6044343	25	6044343CD1	76	6044343CB1	
7503990	26	7503990CD1	77	7503990CB1	
7504655	27	7504655CD1	78	7504655CB1	
7504690	28	7504690CD1	79	7504690CB1	90053710CA2, 90053818CA2
7504720	29	7504720CD1	80	7504720CB1	8079636CA2
7504722	30	7504722CD1	81	7504722CB1	7256019CA2
7504733	31	7504733CD1	82	7504733CB1	1456958CA2, 95134365CA2
7507100	32	7507100CD1	83	7507100CB1	
7503330	33	7503330CD1	84	7503330CB1	
7504519	34	7504519CD1	85	7504519CB1	90023474CA2, 90024742CA2, 90024758CA2
7504705	35	7504705CD1	86	7504705CB1	5486686CA2
7504738	36	7504738CD1	87	7504738CB1	4349337CA2, 559380CA2
7510280	37	7510280CD1	88	7510280CB1	3182544CA2, 5861015CA2, 7354785CA2, 7946775CA2
7503700	38	7503700CD1	89	7503700CB1	
7504685	39	7504685CD1	90	7504685CB1	2072758CA2, 2208146CA2, 2326887CA2, 2635337CA2, 3598548CA2, 4152229CA2, 4156112CA2, 4805920CA2, 95069520CA2, 95069528CA2, 95069552CA2, 95069592CA2, 95069652CA2, 95069668CA2, 95069676CA2, 95069752CA2, 95069776CA2, 95069820CA2, 95069828CA2, 95069852CA2, 95069892CA2
7506844	40	7506844CD1	91	7506844CB1	90123506CA2
7510259	41	7510259CD1	92	7510259CB1	56010433CA2, 56010441CA2, 56010449CA2, 56010541CA2, 56010549CA2, 56010557CA2, 90135770CA2, 90135786CA2
7510444	42	7510444CD1	93	7510444CB1	90001102CA2, 90001110CA2, 90001210CA2
7510494	43	7510494CD1	94	7510494CB1	90017546CA2, 90017614CA2
6486485	44	6486485CD1	95	6486485CB1	95166102CA2, 95166118CA2
7503772	45	7503772CD1	96	7503772CB1	

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
7503773	46	7503773CD1	97	7503773CB1	
7504698	47	7504698CD1	98	7504698CB1	5974456CA2, 6024316CA2, 95118339CA2
7510361	48	7510361CD1	99	7510361CB1	
7507013	49	7507013CD1	100	7507013CB1	90122156CA2, 90122188CA2
7510507	50	7510507CD1	101	7510507CB1	
90106370	51	90106370CD1	102	90106370CB1	90106314CA2, 90106346CA2, 90106370CA2, 90106378CA2, 90106422CA2, 90106470CA2, 90106494CA2, 90106495CA2

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
9	1730616CD1	g23320913	1.0E-100	lung cancer metastasis-related protein [Homo sapiens]
23	7500701CD1	g10444289	3.7E-58	[Homo sapiens] MYG1 (probable metal-dependent hydrolase) homolog
24	7500702CD1	g10444289	3.7E-58	[Homo sapiens] MYG1 homolog
27	7504655CD1	568682|SPUF	2.5E-27	[Homo sapiens] Protein with low similarity to a region of PMBP, which is a putative steroid membrane receptor that is preferentially expressed in the placenta
34	7504519CD1	g24285996	5.0E-22	G-protein coupled receptor GPR114 [Homo sapiens]
35	7504705CD1	336408|MGP	4.2E-11	[Homo sapiens][Structural protein][Extracellular matrix (cuticle and basement membrane); Extracellular (excluding cell wall)] Matrix Gla protein, a vitamin K-dependent calcium-binding component of extracellular matrix that inhibits the calcification of arteries and cartilage; mutation of the corresponding gene causes Keutel syndrome (Chen, L. et al. (1990) Oncogene 5:1391-1395; Munroe, P. B. et al. (1999) Nat. Genet. 21:142-144.)
37	7510280CD1	690518|CKLF1	6.8E-11	[Homo sapiens] Chemokine-like factor 1, a secreted chemoattractant for leucocytes, neutrophils, monocytes and lymphocytes, stimulates inflammatory response and muscle stem cell proliferation; expression is inhibited by IL10 (Han, W. et al. (2001) Biochem. J. 357:127-135.)
45	7503772CD1	g20269129	0.0	MEGF6 [Homo sapiens]
		443807|Y64G10A.7	7.3E-204	[Caenorhabditis elegans] Putative ortholog of D. melanogaster N (Notch) (Zhou, Z. et al. (2001) Cell 104:43-56.)
		716683|MEGF11	7.8E-168	[Homo sapiens] Protein containing five laminin epidermal growth factor (EGF)-like and 15 epidermal growth factor (EGF)-like domains, has low similarity to mouse Notch2
		659407|ced-1	2.9E-151	[Caenorhabditis elegans][Receptor (signalling)][Plasma membrane] Notch family member; putative cell surface phagocytic receptor that recognizes cell corpses in apoptosis (Liu, Q. A. et al. (1998) Cell 93:961-972; Wu, Y. C. et al. (2000) Genes And Development 14:536-548.)
46	7503773CD1	g20269129	0.0	[Rattus norvegicus] MEGF6 (Nakayama, M. et al. (1998) Genomics 51:27-34.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		662841|Egfl3	0.0	[Rattus norvegicus] Multiple epidermal growth factor (EGF)-like domains 6, contains 30 epidermal growth factor-like motifs, putative secreted protein, predicted to bind calcium (Nakayama, M. et al. supra)
		443807|Y64G10A.7	0.0	[Caenorhabditis elegans] Putative ortholog of D. melanogaster N (Notch) (Zhou, Z. et al. (2001) supra.)
		716683|MEGF11	3.3E-192	[Homo sapiens] Protein containing five laminin epidermal growth factor (EGF)-like and 15 epidermal growth factor (EGF)-like domains, has low similarity to mouse Notch2
47	7504698CD1	g7416941	2.0E-29	[Homo sapiens] M5-14
		476049|LOC51300	1.7E-30	[Homo sapiens] M5-14 protein, a putative transmembrane protein that is ubiquitously expressed (Escarceller, M. et al. (2000) DNA Seq. 11:335-338)
48	7510361CD1	g16755850	8.0E-92	[Homo sapiens] parotid secretory protein
		582625|Psp	1.7E-23	[Mus musculus][Extracellular (excluding cell wall)] Parotid secretory protein, a bacteria-binding protein that may function as an innate antimicrobial agent, aberrant expression is observed in nonobese diabetic mice (Robinson, C. P. et al. (1997) Am. J. Physiol. 272:G863-871; Gonzalez, M. J. et al. (2000) Exp. Mol. Pathol. 69:91-101.)
		751684|Psp	2.4E-22	[Rattus norvegicus][Extracellular (excluding cell wall)] Parotid secretory protein (submandibular gland protein A), a major secreted product of the submandibular gland acinar-cell progenitors and parotid glands, may play a role in bacterial binding, expression may be downregulated in diabetes (Mirels, L. et al. (1998) Biochem. J. 330:437-444; Szczepanski, A. et al. (1998) Eur. J. Morphol. Suppl:240-246.)
49	7507013CD1	g2338292	1.5E-12	[Homo sapiens] proline-rich Gla protein 2 (Kulman, J. D. et al. (1997) Proc. Natl. Acad. Sci. U.S.A. 94:9058-9062.)
		337254|PRRG2	1.2E-13	[Homo sapiens][Plasma membrane] Proline-rich Gla (G-carboxyglutamic acid) polypeptide 2, member of the vitamin K-dependent family of proteins (Kulman, J. D. et al. (1997) Proc Natl Acad Sci U S A 94:9058-9062.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
50	7510507CD1	g12856025	1.6E-184	[Mus musculus] (AK016991) WD domain, G-beta repeat containing protein-data source:Pfam, source key:PF00400, evidence:ISS~putative (Carninci, P. et al. (1999) Meth. Enzymol. 303:19-44; Carninci, P. et al. (2000) Genome Res. 10, 1617-1630.)
		742592|DKFZP434C245	7.6E-99	[Homo sapiens] Protein containing WD domains (WD-40 repeat), which may mediate protein-protein interactions
		753721|WDR5	9.4E-37	[Homo sapiens] WD repeat domain 5, contains seven WD domains (WD-40 repeats), which likely mediate protein-protein interactions

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	7506904CD1	120	Signal_cleavage: M1-G37	SPSCAN
			Signal Peptide: M20-G37, M20-T40, M20-C43, M1-G37, M20-S44	HMMER
			Leucine zipper pattern: L5-L26	MOTIFS
			Potential Phosphorylation Sites: S62 T96	MOTIFS
2	7506909CD1	79	Signal_cleavage: M1-G37	SPSCAN
			Signal Peptide: M20-G37, M20-T40, M20-Q43, M1-G37, M20-S42	HMMER
			Leucine zipper pattern: L5-L26	MOTIFS
			Potential Phosphorylation Sites: S64	MOTIFS
3	7507096CD1	270	Signal_cleavage: M1-A21	SPSCAN
			Signal Peptide: M1-S15, M1-A19, M1-A21, M1-G23, M1-S22, M1-W16	HMMER
			Potential Phosphorylation Sites: S22 S31 S41 S222 S240 T155	MOTIFS
4	7507098CD1	225	Signal_cleavage: M1-A60	SPSCAN
			Signal Peptide: M40-S54, M40-A58, M40-A60, M40-G62, M40-S61, M40-W55	HMMER
			MADS-box domain signature and profile: Y87-G169	PROFILESCAN
			Potential Phosphorylation Sites: S61 S70 S80 S171 S181 S220 T190 Y206	MOTIFS
5	7507099CD1	283	Signal_cleavage: M1-A60	SPSCAN
			Signal Peptide: M40-S54, M40-A58, M40-A60, M40-G62, M40-S61, M40-W55	HMMER
			Potential Phosphorylation Sites: S61 S70 S80 S181 S235 S253 T190	MOTIFS
6	7501399CD1	294	Signal Peptide: M1-V16, M1-K18, M1-T21, M1-T23, M1-R25, M1-K22	HMMER
			Potential Glycosylation Sites: N159 N190 N194	MOTIFS
			Potential Phosphorylation Sites: S68 S161 S174 S176 S195 S235 T23 T35 T196 Y162	MOTIFS
7	7504768CD1	40	Signal_cleavage: M1-A20	SPSCAN
			Signal Peptide: M1-S18, M1-A20, M1-C24	HMMER
8	7500757CD1	187	Signal_cleavage: M1-P27	SPSCAN
			Signal Peptide: M1-A24	HMMER
			Immunoglobulin domain: G55-V127	HMMER_PFAM
			Prenyl group binding site (CAAX box):	MOTIFS
			Potential Phosphorylation Sites: S63 S83 S107 S179 T46 T51 T54 T88 Y123	MOTIFS
			Potential Glycosylation Sites: N105	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
9	1730616CD1	261	Signal_cleavage: M1-A29	SPSCAN
			Potential Phosphorylation Sites: S143 S146 S211 S259 T209	MOTIFS
			Potential Glycosylation Sites: N20 N141	MOTIFS
10	190404CD1	506	Signal_cleavage: M1-R20	SPSCAN
			Potential Phosphorylation Sites: S4 S18 S50 S136 S169 S207 S363 S364 T199 T309 T353	MOTIFS
			Potential Glycosylation Sites: N327	MOTIFS
11	7500679CD1	69	Signal_cleavage: M1-P44	SPSCAN
			Signal Peptide: M17-P32, M17-L35, M17-P41, M17-G48, M17-P44, M17-S38	HMMER
			Cell attachment sequence: R50-D52	MOTIFS
			Potential Phosphorylation Sites: T16	MOTIFS
12	7500687CD1	121	Signal_cleavage: M1-S26	SPSCAN
			Signal Peptide: M1-H28, M1-S23, M1-S26	HMMER
			Potential Phosphorylation Sites: S81 S115	MOTIFS
			Potential Glycosylation Sites: N50 N79	MOTIFS
13	7500688CD1	290	Signal_cleavage: M1-G27	SPSCAN
			Signal Peptide: M1-R30, M1-G24, M1-G25, M1-G27	HMMER
			Potential Phosphorylation Sites: S76 S82 S250 S284 T28 T44 T132	MOTIFS
			Potential Glycosylation Sites: N74 N87 N130 N143 N160 N173 N219 N248	MOTIFS
14	7500697CD1	27	Signal Peptide: M1-S19, M1-C20, M1-Q22	HMMER
15	7500709CD1	500	Signal_cleavage: M1-S67	SPSCAN
			Cytosolic domain: M1-R48; Transmembrane domain: A49-P71 Non-cytosolic domain: C72-A500	TMHMMER
			Potential Phosphorylation Sites: S29 S78 S124 S288 S470 T147 T316 T317 T359 T363 T391 T393	MOTIFS
			Potential Glycosylation Sites: N116 N120 N144 N198 N206 N314 N389 N421	MOTIFS
16	7500711CD1	543	Signal_cleavage: M1-S67	SPSCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Cytosolic domain: M1-R48 Transmembrane domain: A49-P71 Non-cytosolic domain: C72-A543	TMHMMER
			Potential Phosphorylation Sites: S29 S78 S124 S279 S435 S437 S513 T138 T307 T308 T350 T354 T382 T384	MOTIFS
			Potential Glycosylation Sites: N116 N120 N135 N189 N197 N305 N380 N464	MOTIFS
17	7500723CD1	72	Signal_cleavage: M1-G28	SPSCAN
			Signal_cleavage: W12-G28, S11-A32, A9-A32, M1-A32, M1-P29	HMMER
			Potential Phosphorylation Sites: S11 T5	MOTIFS
18	7500764CD1	22	Signal Peptide: M1-K22	HMMER
19	7500772CD1	26	Signal Peptide: M1-A16, M1-S19, M1-S20, M1-T22, M1-L24, M1-P18	HMMER
			Potential Phosphorylation Sites: S19 T22	MOTIFS
20	7501350CD1	27	Signal Peptide: M1-P18, M1-A20, M1-G22, M1-A16	HMMER
			Potential Phosphorylation Sites: S19	MOTIFS
21	7506396CD1	253	Signal_cleavage: M1-G54	SPSCAN
			PROTEIN F20D22.3 C47D12.2 PD043239: V21-D160, K205-F248	BLAST_PRODOM
			Potential Phosphorylation Sites: S58 S71 S138 S151 S157 S172 T7 T9 T18 T50	MOTIFS
			Potential Glycosylation Sites: N56 N66 N95 N114	MOTIFS
22	7505917CD1	511	Signal_cleavage: M1-A35	SPSCAN
			PROTEIN CONSERVED PUTATIVE NICOTINATE PHOSPHORIBOSYLTRANSFERASE TRANSFERASE GLYCOSYLTRANSFERASE YUEK CY130.15C 392AA PD008895: F92-E223, E268-E480	BLAST_PRODOM
			CONSERVED HYPOTHETICAL PROTEIN PD183751: P377-L505	BLAST_PRODOM
			PROTEIN CONSERVED PUTATIVE NICOTINATE PHOSPHORIBOSYLTRANSFERASE TRANSFERASE GLYCOSYLTRANSFERASE YUEK CY130.15C PD011757: L16-L80	BLAST_PRODOM
			Potential Phosphorylation Sites: S220 S462 S495 T105 T437	MOTIFS
23	7500701CD1	127	Signal_cleavage: M1-T21	SPSCAN
			Signal Peptide: M1-T21	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			PROTEIN HYDROLASE K08H10.8 BEM2NCB1 INTERGENIC REGION PREDICTED METAL DEPENDENT PD105898: K40-R110	BLAST_PRODOM
			Potential Glycosylation Sites: N51	MOTIFS
24	7500702CD1	137	Signal_cleavage: M1-T21	SPSCAN
			Signal Peptide: M1-T21	HMMER
			PROTEIN HYDROLASE K08H10.8 BEM2NCB1 INTERGENIC REGION PREDICTED METAL DEPENDENT PD105898: K40-R110	BLAST_PRODOM
			Potential Glycosylation Sites: N51	MOTIFS
25	6044343CD1	207	Signal_cleavage: M1-G46	SPSCAN
			Potential Phosphorylation Sites: S153 S181 T80 T84 T92 T98 T102 T106 T114 T146 T150	MOTIFS
			Potential Glycosylation Sites: N174	MOTIFS
26	7503990CD1	1008	signal_cleavage: M1-A34	SPSCAN
			Potential Phosphorylation Sites: S95 S168 S245 S276 S337 S375 S407 S411 S434 S457 S501 S531 S548 S564 S580 S625 S670 S680 S684 S761 S792 S800 S804 S848 S850 S882 S891 S924 S971 T49 T68 T162 T166 T347 T362 T419 T588 T731 T777 T778 T912 T967 T975 T1006	MOTIFS
			Potential Glycosylation Sites: N47 N142 N172 N207 N225 N226 N230 N586	MOTIFS
27	7504655CD1	70	Signal_cleavage: M1-G25	SPSCAN
			Signal Peptide: M1-A31	HMMER
			Potential Phosphorylation Sites: S60 T28 T34 T48	MOTIFS
28	7504690CD1	142	Signal_cleavage: M1-L19	SPSCAN
			Potential Phosphorylation Sites: S83 S110 T88 T92 T105 T136	MOTIFS
			Signal Peptide: M1-A17, M1-L19, M1-Q21, M1-S24	HMMER
29	7504720CD1	43	Signal_cleavage: M1-G20	SPSCAN
			Signal Peptide: M1-G20, M1-T22, M1-K24, M1-R28	HMMER
			Potential Phosphorylation Sites: S40 T22 T26 T33	MOTIFS
30	7504722CD1	64	Signal_cleavage: M1-R33	SPSCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Potential Phosphorylation Sites: S42	MOTIFS
31	7504733CD1	23	Signal Peptide: M1-S15	HMMER
			Signal Peptide: M1-V18, M1-P20, M1-L16, M1-P20	HMMER
32	7507100CD1	349	Signal_cleavage: M1-A60	SPSCAN
			Signal Peptide: M40-S54, M40-A58, M40-A60, M40-G62, M40-S61, M40-W55	HMMER
			Potential Phosphorylation Sites: S61 S70 S80 S181 S240 S289 T190 T265 T344 Y206	MOTIFS
			Potential Glycosylation Sites: N238 N288	MOTIFS
33	7503330CD1	65	Signal_cleavage: M1-A27	SPSCAN
			Signal Peptide: M9-A27, M9-S28, M9-A31, M9-L33, M1-S28, M1-A31	HMMER
			Potential Phosphorylation Sites: S5	MOTIFS
34	7504519CD1	64	Signal_cleavage: M1-A19	SPSCAN
			Signal Peptide: M1-A19, M1-T21, M1-T23, M1-E25, M1-E22	HMMER
			Potential Phosphorylation Sites: S29 S45 S46 T23	MOTIFS
			Potential Glycosylation Sites: N18	MOTIFS
35	7504705CD1	94	Signal_cleavage: M1-C19	SPSCAN
			Signal Peptide: M1-V16, M1-C19, M1-E21, M1-S22, M1-E24, M1-L18	HMMER
			Potential Phosphorylation Sites: S87	MOTIFS
			Potential Glycosylation Sites: N78	MOTIFS
36	7504738CD1	52	Signal_cleavage: M1-S23	SPSCAN
			Signal Peptide: M1-A16, M1-C18, M1-P20, M1-S23, M1-A24, M1-P25	HMMER
			Cytosolic domain: M1-T6 Transmembrane domain: V7-A24 Non-cytosolic domain: P25-N52	TMHMMER
37	7510280CD1	33	Potential Phosphorylation Sites: S16	MOTIFS
38	7503700CD1	242	Signal_cleavage: M1-A20	SPSCAN
			Signal Peptide: M1-A19, M1-S21, M1-A29, M1-A20, M1-G24	HMMER
			Potential Phosphorylation Sites: S21 S174 T46 T176 T203	MOTIFS
			Potential Glycosylation Sites: N220	MOTIFS
39	7504685CD1	47	Signal_cleavage: M1-A18	SPSCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Signal Peptide: M1-A18, M1-R20	HMMER
			Cytosolic domain: Q24-Q47 Transmembrane domain: T4-L23 Non-cytosolic domain: M1-S3	TMHMMER
40	7506844CD1	53	Signal_cleavage: M1-E24	SPSCAN
			Signal Peptide: M1-G21, M1-E24	HMMER
41	7510259CD1	95	Signal_cleavage: M1-G46	SPSCAN
			Cytosolic domain: M1-R29 Transmembrane domain: L30-V52 Non-cytosolic domain: V53-R95	TMHMMER
			Guanylate cyclases signature: L35-R93	PROFILESCAN
			Potential Phosphorylation Sites: S9 S14 S27 S64 S80	MOTIFS
42	7510444CD1	57	Signal_cleavage: M1-L19	SPSCAN
			Signal Peptide: M1-L19, M1-S20, M1-A22, M1-K24, M1-I25, M1-T27	HMMER
			Cytosolic domain: M1-V6 Transmembrane domain: L7-S29 Non-cytosolic domain: T30-D57	TMHMMER
			Potential Phosphorylation Sites: T30 T49	MOTIFS
43	7510494CD1	67	Signal_cleavage: M1-A21	SPSCAN
			Signal Peptide: M1-S17, M1-C18, M1-A21, M1-H23, M1-A24, M1-L30	HMMER
			Potential Phosphorylation Sites: S28 S51	MOTIFS
			Potential Glycosylation Sites: N48	MOTIFS
44	6486485CD1	311	Signal Peptide: M1-A26	HMMER
			Signal Peptide: M1-S31, M9-S31	HMMER
			Immunoglobulin domain: T97-L194	HMMER_SMART
			Ig superfamily from SCOP: T93-S197	HMMER_INCY
			Cytosolic domain: R232-G311 Transmembrane domain: L209-T231 Non-cytosolic domain: M1-P208	TMHMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Potential Phosphorylation Sites: S36 S67 S147 S197 S250 S297 T166 T179 T186 T231	MOTIFS
			Potential Glycosylation Sites: N99 N168 N303	MOTIFS
45	7503772CD1	1097	signal_cleavage: M1-G26	SPSCAN
			Signal Peptide: M1-A22, M1-V24, M1-Y29, M1-G26	HMMER
			EGF-like domain: C60-C95, C101-C137, C143-C178, C184-C219, C230-C265, C271-C305, C311-C346, C415-C446, C459-C489, C502-C532, C536-C577, C590-C622, C635-C664, C677-C708, C721-C751, C764-C795, C808-C838, C851-C881, C894-C924, C937-C967, C980-C1010, C1023-C1053	HMMER_PFAM
			Epidermal growth factor-like domain: E59-L96, S100-V138, P142-E179, E183-Y220, S229-I266, D270-E306, E310-S347, D414-N447, N458-E490, H501-H533, T535-Q578, E580-G623, N634-E665, G676-Q709, S720-Q752, D763-E796, G807-E839, D850-A882, N893-L925, E927-Q968, A979-E1011, G1022-N1054	HMMER_SMART
			Calcium-binding EGF-like domain: D56-L96, A97-V138, C143-E179, D180-Y220, V227-I266, D267-G304, D307-S347, C459-E490, Q618-E665, C681-Q709, C721-Q752, C894-L925, E927-Q968, C980-E1011	HMMER_SMART
			Laminin-type epidermal growth factor-like domain: C463-C502, C506-C545, C549-C590, C594-C635, C639-C677, C681-C721, C725-C764, C768-C808, C812-C851, C855-C894, C898-C937, G942-C980, C984-C1023, C1027-C1069	HMMER_SMART
			Calcium-binding EGF-like domain proteins pattern proteins BL01187: C265-C276, C71-F86	BLIMPS_BLOCKS
			Type III EGF-like signature PR00011: C471-C489, C690-C708	BLIMPS_PRINTS
			Thrombomodulin signature PR00907: C237-P253, L258-C281, G286-C311	BLIMPS_PRINTS
			MEGF6 GLYCOPROTEIN EGF-LIKE DOMAIN PD169326: L349-D414 PD165309: N507-C536	BLAST_PRODOM
			SURFACE ANTIGEN PROTEIN PRECURSOR SIGNAL REPEAT MEMBRANE GPI ANCHOR 156G 168G PD001714: T470-C924	BLAST_PRODOM
			PROTEIN TRANSCRIPTIONAL REPEAT TRANSCRIPTION REGULATION DNA-BINDING NUCLEAR SHUTTLE CRAFT PUTATIVE PD014613: C406-C881	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			EGF DM00003|P98163|1373-1460: S229-V308, C277-E350, C143-E223 DM00003|P35556|2219-2292: S274-C346	BLAST_DOMO
			EGF-LIKE DOMAIN DM00864|I55476|159-241: N234-C311, C277-D355, R146-E223	BLAST_DOMO
			SUSHI REPEAT DM04887|P16581|1-609: C459-A882	BLAST_DOMO
			Aspartic acid and asparagine hydroxylation site: C71-C82, C195-C206, C281-C292, C322-C333	MOTIFS
			EGF-like domain signature 1: C435-C446, C478-C489, C521-C532, C566-C577, C653-C664, C697-C708, C740-C751, C784-C795, C827-C838, C870-C881, C913-C924, C956-C967, C999-C1010, C1042-C1053	MOTIFS
			EGF-like domain signature 2: C80-C95, C122-C137, C163-C178, C204-C219, C250-C265, C290-C303, C331-C346, C435-C446, C566-C577, C697-C708, C740-C751, C784-C795, C827-C838, C913-C924, C956-C967, C999-C1010, C1042-C1053	MOTIFS
			Calcium-binding EGF-like domain pattern signature: D56-C80, D180-C204, D267-C290, D307-C331	MOTIFS
			Potential Phosphorylation Sites: S49 S201 S313 S347 S567 S1008 T264 T399 T401 T418 T487 T660 T836 T953	MOTIFS
			Potential Glycosylation Sites: N147 N447 N458 N634 N769 N856 N867 N893 N1054	MOTIFS
46	7503773CD1	1350	signal_cleavage: M1-G26	SPSCAN
			Signal Peptide: M1-A22, M1-V24, M1-Y29, M1-G26	HMMER
			EGF-like domain: C60-C95, C101-C137, C143-C178, C184-C219, C230-C265, C271-C305, C311-C346, C415-C446, C459-C489, C502-C532, C536-C577, C590-C622, C635-C664, C677-C708, C721-C751, C764-C795, C808-C838, C851-C881, C894-C924, C937-C967, C980-C1010, C1014-C1054, C1067-C1097, C1110-C1140, C1153-C1183, C1196-C1226, C1239-C1269, C1282-C1312	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Epidermal growth factor-like domain: E59-L96, S100-V138, P142-E179, E183-Y220, S229-I266, D270-E306, E310-S347, D414-N447, N458-E490, H501-H533, T535-Q578, E580-G623, N634-E665, G676-Q709, S720-Q752, D763-E796, G807-E839, D850-A882, N893-L925, E927-Q968, A979-E1011, G1022-Q1055, G1066-N1098, N1109-E1141, G1152-E1184, A1195-E1227, P1229-E1270, G1281-N1313	HMMER_SMART
			Calcium-binding EGF-like domain: D56-L96, A97-V138, C143-E179, D180-Y220, V227-I266, D267-G304, D307-S347, C459-E490, Q618-E665, C681-Q709, C721-Q752, C894-L925, E927-Q968, C980-E1011, C1153-E1184, E1199-E1227	HMMER_SMART
			Laminin-type epidermal growth factor-like domain: C463-C502, C506-C545, C549-C590, C594-C635, C639-C677, C681-C721, C725-C764, C768-C808, C812-C851, C855-C894, C898-C937, G942-C980, C984-C1023, C1027-C1067, C1071-C1110, C1114-C1153, C1157-C1196, C1200-C1239, C1243-C1282, C1286-N1313	HMMER_SMART
			Calcium-binding EGF-like domain proteins pattern proteins BL01187: C71-F86	BLIMPS_BLOCKS
			Type III EGF-like signature PR00011: C471-C489, C690-C708	BLIMPS_PRINTS
			Thrombomodulin signature PR00907: C237-P253, L258-C281, G286-C311	BLIMPS_PRINTS
			MEGF6 GLYCOPROTEIN EGF-LIKE DOMAIN PD169326: L349-D414 PD165309: N507-C536	BLAST_PRODOM
			SURFACE ANTIGEN PROTEIN PRECURSOR SIGNAL REPEAT MEMBRANE GPI ANCHOR 156G 168G PD001714: H774-C1215	BLAST_PRODOM
			PROTEIN TRANSCRIPTIONAL REPEAT TRANSCRIPTION REGULATION DNA-BINDING NUCLEAR SHUTTLE CRAFT PUTATIVE PD014613: C406-C881, G802-C1312	BLAST_PRODOM
			EGF DM00003|P98163|1373-1460: S229-V308, C277-E350, C143-E223 DM00003|P35556|2219-2292: S274-C346 DM00003|A57278|2213-2286: S274-C346	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			EGF-LIKE DOMAIN DM00864|I55476|159-241: N234-C311, C277-D355, R146-E223	BLAST_DOMO
			Aspartic acid and asparagine hydroxylation site: C71-C82, C195-C206, C281-C292, C322-C333	MOTIFS
			EGF-like domain signature 1: C435-C446, C478-C489, C521-C532, C566-C577, C653-C664, C697-C708, C740-C751, C784-C795, C827-C838, C870-C881, C913-C924, C956-C967, C999-C1010, C1043-C1054, C1086-C1097, C1129-C1140, C1172-C1183, C1215-C1226, C1258-C1269, C1301-C1312	MOTIFS
			EGF-like domain signature 2: C80-C95, C122-C137, C163-C178, C204-C219, C250-C265, C290-C303, C331-C346, C435-C446, C566-C577, C697-C708, C740-C751, C784-C795, C827-C838, C913-C924, C956-C967, C999-C1010, C1043-C1054, C1086-C1097, C1172-C1183, C1215-C1226, C1258-C1269	MOTIFS
			Calcium-binding EGF-like domain pattern signature: D56-C80, D180-C204, D267-C290, D307-C331	MOTIFS
			Potential Phosphorylation Sites: S49 S201 S313 S347 S567 S1008 S1018 S1158 S1206 S1255 S1277 S1337 T264 T399 T401 T418 T487 T660 T836 T953 T1179 T1214 T1341	MOTIFS
			Potential Glycosylation Sites: N147 N447 N458 N634 N769 N856 N867 N893 N1098 N1109 N1169 N1204 N1205	MOTIFS
47	7504698CD1	71	signal_cleavage: M1-A24	SPSCAN
			Signal Peptide: M1-A25, M1-A30, M1-A24	HMMER
			Vitamin K-dependent carboxylation domain: V3-F71	PROFILESCAN
			ATP/GTP-binding site motif A (P-loop): G63-S70	MOTIFS
			Potential Phosphorylation Sites: S53	MOTIFS
48	7510361CD1	220	signal_cleavage: M1-S18	SPSCAN
			Signal Peptide: M1-T15, M1-E19, M1-S20	HMMER
			PROTEIN SALIVARY PRECURSOR GLAND SIGNAL BSP30 PAROTID SECRETORY PSP SUBMANDIBULAR PD011295: M1-I192	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			PARSOTID SECRETORY PROTEIN DM04779|P07743|12-234: G12-L186 DM04779|B42337|12-235: G12-L186	BLAST_DOMO
			Leucine zipper pattern: L44-L65	MOTIFS
			Potential Phosphorylation Sites: S20 S184 T50 T95	MOTIFS
			Potential Glycosylation Sites: N124 N132	MOTIFS
49	7507013CD1	39	Signal_cleavage: M1-D19	SPSCAN
			Signal Peptide: M1-D19, M1-P22, M1-E25, M1-S23, M1-T20	HMMER
			PEP-utilizing enzymes signatures: M1-V38	PROFILESCAN
			PROLINERICH GLA PROTEIN 2 PD059428: M1-G33	BLAST_PRODOM
			Potential Phosphorylation Sites: S21 T16 T26	MOTIFS
50	7510507CD1	451	Signal_cleavage: M1-A53	SPSCAN
			WD domain, G-beta repeat: F136-D172, L220-D256, K94-S130, R52-I88, L262-R298, L10-N46, C178-D214	HMMER_PFAM
			WD repeats: D7-N46, P49-I88, K91-S130, R133-D172, N175-D214, V217-D256, E259-R298	HMMER_SMART
			one copy of WD repeat: L10-N46, R52-I88, G93-S130, R135-D172, T174-D214, V217-D256, G260-R298	HMMER_SMART
			Trp-Asp (WD) repeat proteins proteins BL00678: S161-W171	BLIMPS_BLOCKS
			Trp-Asp (WD-40) repeats signature: C149-G198, D249-F278, V192-F236, T64-A112, T22-S79, S107-P154	PROFILESCAN
			Beta G-protein (transducin) signature PR00319: I159-T173, P196-W213	BLIMPS_PRINTS
			G-protein beta WD-40 repeat signature PR00320: I159-T173	BLIMPS_PRINTS
			PROTEIN REPEAT WD TRPASP REPEATS CONTAINING CHROMOSOME NUCLEAR FACTOR I PD000061: S203-L257, S96-W129	BLAST_PRODOM
			Trp-Asp (WD) repeats signature: L33-F47, I159-T173, I201-V215	MOTIFS
			Potential Phosphorylation Sites: S3 S125 S161 S349 S393 S420 T22 T83 T119 T167 T174 T209 T251 T299 T310 T352 T378 T379 T383 T427 Y363	MOTIFS
			Potential Glycosylation Sites: N181	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
51	90106370CD1	224	Signal_cleavage: M1-A53	SPSCAN
			Signal Peptide: M35-A53	HMMER
			Cytosolic domain: E54-F224 Transmembrane domain: C31-A53 Non-cytosolic domain: M1-P30	TMHMMER
			Potential Phosphorylation Sites: S9 S66 S168 T87 T104 T141 Y96	MOTIFS

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
52/ 7506904CB1/ 1480	1-122, 1-220, 1-540, 1-560, 1-599, 1-608, 1-640, 1-641, 4-1473, 40-548, 40-603, 40-613, 44-443, 94-722, 94-774, 94-803, 94-836, 94-865, 95-480, 95-839, 95-865, 96-590, 96-751, 96-768, 102-865, 104-865, 113-865, 154-865, 183-857, 204-865, 223-865, 245-586, 257-586, 276-611, 293-643, 353-464, 359-643, 407-530, 506-632, 638-865, 645-865, 757-1197, 758-1007, 776-1432, 815-1048, 819-1026, 826-1416, 888-1341, 980-1103, 1052-1341, 1070-1341, 1092-1341, 1102-1480, 1104-1480, 1126-1339, 1173-1364, 1260-1341
53/ 7506909CB1/ 1168	1-122, 1-220, 1-280, 4-1168, 40-319, 75-261, 94-341, 94-560, 94-563, 95-356, 95-563, 105-563, 170-341, 322-563, 338-563, 339-562, 339-563, 343-563, 455-895, 456-705, 474-1130, 513-746, 517-724, 524-1114, 586-1039, 668-1165, 678-801, 750-1039, 768-1039, 790-1039, 800-1168, 824-1037, 871-1062, 958-1039
54/ 7507096CB1/ 2229	1-2229, 165-521, 165-536, 171-543, 174-325, 174-393, 178-541, 178-956, 178-1105, 296-380, 415-1271, 476-1272, 558-1093, 558-1096, 558-1266, 561-1266, 570-981, 588-1271, 608-1271, 657-1364, 678-803, 701-1266, 770-1266, 793-1007, 799-1317, 913-1742, 993-1153, 1012-1318, 1012-1330, 1109-1624, 1135-1559, 1140-1748, 1164-1455, 1232-1443, 1232-1483, 1232-1624, 1232-1668, 1232-1669, 1232-1687, 1232-1751, 1232-1891, 1236-1507, 1252-1621, 1284-1846, 1352-1994, 1371-1694, 1371-1992, 1431-1698, 1454-2017, 1454-2028, 1460-2029, 1476-1796, 1478-1983, 1478-1992, 1488-1767, 1491-2225, 1546-2009, 1565-1749, 1573-1820, 1573-1983, 1606-1960, 1640-1953, 1666-2226, 1686-1950, 1690-2041, 1690-2042, 1698-1897, 1728-2053, 1733-1968, 1741-2105, 1745-2037, 1755-2086, 1799-2026, 1801-1980, 1810-1998, 1842-2041, 1843-2041, 1900-2044, 1946-2225, 2065-2229
55/ 7507098CB1/ 2374	1-2374, 165-521, 165-536, 171-542, 174-325, 174-393, 178-541, 178-706, 178-939, 296-380, 703-1238, 703-1241, 703-1411, 703-1417, 706-1411, 706-1416, 715-1126, 733-1416, 753-1416, 802-1358, 823-948, 846-1411, 915-1411, 938-1152, 944-1212, 1058-1887, 1138-1298, 1157-1340, 1157-1406, 1254-1769, 1280-1704, 1285-1893, 1309-1600, 1377-1515, 1377-1628, 1377-1769, 1377-1813, 1377-1814, 1377-1832, 1377-1896, 1377-2036, 1381-1652, 1397-1766, 1429-1991, 1516-1839, 1516-2137, 1576-1843, 1599-2162, 1599-2173, 1605-2174, 1621-1941, 1623-2128, 1623-2137, 1633-1912, 1691-2154, 1710-1894, 1718-1919, 1718-1965, 1751-2105, 1785-2098, 1811-2184, 1831-2095, 1835-2142, 1843-2042, 1844-2370, 1872-2139, 1873-2142, 1878-2113, 1886-2250, 1890-2182, 1900-2231, 1944-2171, 1946-2125, 1955-2143, 1987-2185, 1988-2185, 2045-2189, 2091-2175, 2210-2374
56/ 7507099CB1/ 2151	1-2151, 165-521, 165-536, 171-543, 174-325, 174-393, 178-541, 178-661, 296-380, 560-1193, 692-1188, 715-929, 721-960, 835-1664, 915-1075, 934-1117, 934-1183, 1031-1546, 1057-1365, 1062-1670, 1086-1377, 1154-1292, 1154-1405, 1154-1464, 1154-1590, 1154-1591, 1154-1609, 1154-1673, 1154-1813, 1158-1429, 1174-1543, 1206-1768, 1274-1916, 1293-1616, 1293-1914, 1353-1620, 1376-1939, 1376-1950, 1382-1951, 1398-1718, 1400-1905, 1400-1914, 1410-1689, 1413-2147, 1468-1931, 1487-1671, 1495-1696, 1495-1742, 1528-1882, 1562-1875, 1588-1961, 1608-1872, 1612-1963, 1612-1964, 1620-1819, 1650-1975, 1655-1890, 1663-2027, 1667-1959, 1677-2008, 1721-1948, 1723-1902, 1732-1920, 1764-1963, 1765-1963, 1822-1966, 1868-1979, 1987-2151

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
57/ 7501399CB1/ 3435	1-492, 1-687, 1-3362, 26-637, 34-606, 46-591, 98-722, 98-775, 110-594, 164-649, 316-834, 342-596, 350-903, 368-911, 491-770, 540-717, 540-785, 540-1048, 540-1068, 540-1091, 540-1311, 540-1341, 787-1371, 789-3362, 903-1615, 913-1359, 921-1615, 930-1516, 941-1599, 941-1615, 960-1615, 964-1615, 965-1622, 994-1615, 1003-1611, 1009-1428, 1009-1615, 1011-1615, 1012-1082, 1021-1615, 1022-1615, 1024-1615, 1026-1615, 1031-1601, 1032-1615, 1032-1619, 1036-1671, 1037-1615, 1043-1615, 1049-1675, 1059-1728, 1062-1615, 1071-1509, 1071-1615, 1097-1728, 1098-1615, 1104-1733, 1114-1743, 1139-1615, 1150-1615, 1152-2015, 1167-1430, 1170-1640, 1174-1615, 1200-1823, 1203-1744, 1228-1774, 1244-1747, 1254-1828, 1272-1615, 1281-1614, 1290-1771, 1298-1783, 1302-1988, 1303-1579, 1306-1874, 1307-1615, 1312-1615, 1326-1928, 1328-1915, 1346-1825, 1395-1797, 1406-1886, 1429-1610, 1483-1749, 1531-2090, 1541-2138, 1541-2173, 1541-2193, 1554-2069, 1620-1803, 1629-1879,
	1629-2034, 1629-2173, 1629-2265, 1629-2293, 1636-2200, 1661-2106, 1670-2244, 1697-2333, 1699-1797, 1712-2339, 1718-1907, 1718-2292, 1719-2276, 1726-2310, 1764-2569, 1780-2398, 1801-1885, 1817-2332, 1824-2045, 1824-2298, 1825-2447, 1846-2195, 1849-2083, 1851-2085, 1851-2102, 1857-2459, 1881-2496, 1921-2429, 1929-2478, 1929-2517, 1944-2589, 1950-2225, 1950-2531, 1983-2558, 1987-2658, 1994-2541, 1996-2368, 2020-2268, 2049-2633, 2053-2314, 2057-2267, 2064-2658, 2068-2634, 2073-2690, 2093-2786, 2095-2628, 2112-2703, 2116-2962, 2119-2404, 2119-2781, 2138-2743, 2148-2394, 2188-2487, 2212-2461, 2218-2768, 2260-2499, 2264-2521, 2265-2491, 2282-3075, 2304-2518, 2329-2750, 2352-2939, 2355-3005, 2365-2947, 2378-2988, 2383-3008, 2411-3050, 2428-2507, 2438-2716, 2442-3049, 2449-2731, 2455-2703, 2476-3133, 2477-2725, 2494-3172, 2498-2932, 2506-3083, 2508-3127, 2528-2816, 2528-3101, 2532-3105, 2535-2802, 2546-3143, 2556-3011, 2557-2854,
	2562-2839, 2570-2803, 2580-2863, 2591-3137, 2597-2853, 2599-3069, 2600-3171, 2601-2877, 2608-3155, 2620-3148, 2625-2831, 2626-2889, 2642-3223, 2645-3096, 2649-3243, 2653-2905, 2653-2935, 2658-2960, 2663-2904, 2683-2984, 2685-3258, 2701-2928, 2733-3358, 2735-2982, 2744-3067, 2749-3271, 2755-3339, 2756-3374, 2759-3016, 2760-3007, 2765-3034, 2778-3072, 2780-2911, 2795-3105, 2807-3072, 2817-3076, 2830-3285, 2835-3113, 2849-3119, 2854-3423, 2858-2987, 2870-3431, 2885-3124, 2904-3167, 2904-3168, 2905-3138, 2905-3435, 2914-3191, 2914-3240, 2918-3188, 2930-3212, 2930-3238, 2940-3263, 2941-3098, 2944-3197, 2956-3270, 2964-3212, 2968-3241, 2992-3216, 2998-3215, 3011-3294, 3026-3435, 3031-3295, 3043-3285, 3052-3288, 3055-3357, 3056-3321, 3065-3312, 3084-3335, 3097-3334, 3097-3367, 3101-3377, 3103-3289, 3123-3381, 3125-3409, 3129-3362
58/ 7504768CB1/ 751	1-279, 2-273, 7-751, 82-220, 127-365, 127-392, 127-413, 132-400, 132-406, 132-407, 134-392, 140-376, 144-400, 153-305, 166-302, 169-404, 181-256, 181-461, 189-366, 191-460, 198-467, 200-470, 214-483, 215-481, 216-406, 216-455, 216-463, 227-369, 235-417, 238-409, 238-447, 281-550, 306-550

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
59/ 7500757CB1/ 1908	1-422, 1-1908, 6-200, 34-462, 75-383, 75-436, 75-491, 419-681, 504-707, 504-1166, 508-1103, 513-1061, 533-664, 564-811, 577-827, 654-1204, 696-919, 696-1300, 755-1047, 756-1170, 756-1376, 780-999, 805-946, 835-1217, 837-1101, 845-1432, 867-1149, 889-1140, 889-1301, 889-1460, 934-1211, 943-1558, 949-1237, 979-1410, 985-1521, 988-1220, 988-1508, 992-1854, 999-1263, 1002-1596, 1011-1237, 1017-1133, 1021-1294, 1036-1711, 1044-1315, 1044-1327, 1059-1612, 1064-1375, 1081-1364, 1081-1771, 1119-1546, 1126-1368, 1131-1423, 1137-1611, 1141-1404, 1146-1415, 1162-1352, 1162-1657, 1176-1865, 1201-1662, 1201-1845, 1204-1876, 1205-1748, 1212-1713, 1233-1810, 1234-1878, 1235-1435, 1238-1481, 1242-1853, 1245-1497, 1247-1490, 1251-1451, 1251-1812, 1253-1513, 1255-1878, 1266-1856, 1279-1823, 1289-1393, 1289-1867, 1290-1874, 1302-1529, 1303-1566, 1308-1872, 1308-1888, 1311-1881, 1312-1597, 1320-1590, 1331-1593, 1331-1885, 1333-1620, 1334-1908, 1343-1879, 1348-1865,
	1359-1888, 1377-1698, 1380-1869, 1384-1881, 1393-1881, 1412-1852, 1422-1864, 1429-1872, 1431-1873, 1436-1856, 1437-1679, 1437-1822, 1439-1715, 1442-1872, 1443-1870, 1447-1880, 1448-1873, 1460-1866, 1461-1866, 1461-1867, 1480-1901, 1490-1885, 1491-1884, 1493-1866, 1501-1874, 1503-1867, 1514-1888, 1521-1888, 1524-1749, 1524-1857, 1524-1878, 1540-1880, 1550-1866, 1561-1867, 1567-1866, 1582-1880, 1623-1888, 1629-1866, 1642-1872, 1652-1866, 1669-1880, 1688-1866, 1692-1870, 1717-1864, 1751-1880, 1753-1866, 1765-1872, 1776-1881, 1800-1871
60/ 1730616CB1/ 1148	1-166, 66-194, 66-279, 67-237, 75-685, 83-321, 83-347, 85-313, 85-314, 85-316, 85-341, 85-393, 85-394, 85-558, 85-727, 85-728, 85-821, 86-293, 86-336, 86-545, 87-338, 91-340, 93-201, 93-340, 93-343, 93-351, 93-381, 94-630, 94-666, 94-683, 95-327, 95-484, 95-846, 96-414, 101-512, 101-708, 106-366, 107-666, 110-626, 111-289, 111-378, 114-709, 115-397, 145-371, 145-377, 178-406, 185-640, 186-851, 219-732, 227-633, 233-732, 240-637, 258-712, 264-708, 277-704, 278-712, 280-543, 295-888, 300-712, 302-704, 304-704, 318-480, 318-561, 318-994, 319-575, 334-577, 334-582, 334-790, 349-593, 357-860, 362-925, 365-616, 375-704, 376-594, 378-708, 379-704, 394-648, 394-711, 406-704, 409-715, 424-705, 435-670, 445-704, 445-718, 445-747, 452-721, 486-683, 486-731, 486-759, 486-761, 486-779, 498-586, 559-1087, 595-1085, 603-871, 624-1148, 627-933, 649-712, 681-1017, 708-1132, 712-1148, 713-1132, 715-953, 716-956, 720-1132, 721-947, 721-1126, 721-1132, 722-1115, 727-1129, 727-1132, 727-1144, 730-
	1148, 731-966, 731-1104, 731-1138, 734-1127, 734-1132, 743-1132, 743-1148, 744-1132, 745-1148, 751-1132, 751-1148, 755-1132, 757-1020, 759-1057, 760-1148, 761-1132, 765-1132, 766-1148, 825-1131, 837-1132, 839-1112, 839-1132, 839-1147, 840-1132, 844-941, 847-1132, 863-1131, 873-1132, 882-1144, 922-1132, 939-1128, 940-1132, 965-1128, 965-1132, 966-1148, 1022-1132, 1091-1144
61/ 190404CB1/ 1776	1-196, 16-446, 78-763, 108-524, 145-502, 242-763, 267-519, 331-411, 446-1043, 735-980, 791-1012, 957-1502, 1134-1776, 1418-1474

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
62/ 7500679CB1/ 1765	1-1765, 136-372, 214-428, 214-436, 214-876, 215-797, 222-513, 236-488, 240-518, 242-484, 246-492, 246-527, 257-821, 268-978, 273-886, 278-392, 283-888, 284-543, 286-820, 289-573, 289-898, 290-619, 293-591, 295-959, 309-907, 310-741, 316-534, 325-679, 332-713, 336-610, 339-713, 355-623, 362-970, 367-605, 367-636, 369-969, 372-637, 375-871, 379-652, 382-947, 392-840, 393-629, 394-912, 395-605, 395-628, 395-1044, 398-665, 400-654, 400-683, 404-876, 409-658, 425-683, 431-754, 433-717, 434-697, 438-1016, 440-877, 445-631, 448-869, 449-709, 450-699, 453-860, 454-1033, 455-695, 458-742, 473-782, 487-1035, 488-741, 501-933, 508-1138, 509-613, 510-770, 513-746, 520-1060, 521-730, 521-947, 526-679, 535-1190, 535-1260, 541-824, 543-758, 543-765, 543-1042, 543-1044, 547-1162, 548-835, 554-797, 560-834, 561-747, 561-1039, 562-858, 563-1059, 564-777, 564-1077, 565-858, 567-1106, 570-793, 571-837, 572-1167, 573-1220, 573-1239, 578-855, 585-858, 587-877, 587-1233, 588-1206, 591-1117, 591-
	1339, 591-1367, 595-874, 597-883, 598-791, 603-893, 631-1080, 634-912, 638-1258, 642-880, 644-1244, 648-903, 648-908, 653-924, 673-750, 673-1300, 678-1254, 684-1175, 693-912, 702-1296, 703-1295, 707-960, 707-1376, 712-920, 712-927, 717-1368, 719-1254, 725-1195, 726-1325, 728-972, 744-1000, 746-873, 752-1045, 758-948, 759-953, 759-967, 759-968, 759-1360, 759-1446, 763-1260, 764-1381, 773-1392, 778-1025, 781-1034, 781-1073, 782-1053, 783-1388, 789-1063, 793-1267, 797-1253, 801-1347, 804-1396, 815-1347, 817-1013, 819-1371, 823-1046, 823-1486, 823-1504, 824-1253, 825-985, 825-1044, 839-1090, 839-1104, 845-934, 845-1090, 849-1112, 849-1450, 855-1103, 862-1163, 864-1202, 867-1161, 870-1136, 876-1197, 883-1137, 885-1102, 885-1121, 885-1628, 886-1132, 887-1083, 892-1070, 892-1131, 893-1252, 896-1117, 918-1693, 920-1615, 922-1172, 924-1160, 925-1175, 928-1058, 928-1198, 929-1133, 930-1467, 931-1421, 932-1215, 932-1220, 936-1205, 936-1340, 936-1740, 949-1175, 949-1712, 958-1211,
	959-1250, 959-1260, 961-1207, 961-1300, 964-1127, 964-1221, 964-1246, 964-1399, 964-1598, 966-1687, 971-1354, 972-1472, 973-1522, 975-1338, 980-1230, 984-1752, 987-1628, 989-1709, 990-1542, 990-1700, 997-1373, 999-1534, 1001-1542, 1007-1235, 1010-1225, 1010-1703, 1012-1211, 1012-1305, 1012-1526, 1012-1542, 1015-1750, 1023-1606, 1030-1567, 1030-1608, 1031-1295, 1035-1316, 1036-1292, 1036-1542, 1036-1546, 1037-1241, 1038-1297, 1043-1700, 1045-1280, 1047-1753, 1048-1331, 1053-1476, 1054-1318, 1059-1326, 1069-1459, 1071-1542, 1072-1360, 1075-1542, 1076-1326, 1077-1340, 1077-1448, 1081-1681, 1082-1720, 1083-1373, 1086-1301, 1089-1332, 1090-1540, 1091-1540, 1098-1710, 1100-1353, 1102-1539, 1105-1345, 1105-1709, 1114-1709, 1119-1361, 1121-1239, 1121-1701, 1123-1375, 1133-1759, 1135-1748, 1138-1616, 1149-1654, 1152-1540, 1158-1734, 1161-1403, 1173-1542, 1174-1540, 1174-1542, 1175-1303, 1185-1449, 1185-1492, 1190-1448, 1193-1458, 1193-1475, 1193-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1478, 1196-1432, 1196-1762, 1198-1765, 1199-1692, 1200-1399, 1201-1459, 1202-1709, 1203-1469, 1205-1542, 1205-1762, 1211-1455, 1211-1503, 1216-1497, 1218-1755, 1221-1395, 1221-1692, 1223-1481, 1230-1496, 1238-1462, 1241-1489, 1241-1495, 1243-1522, 1261-1765, 1262-1459, 1262-1765, 1263-1380, 1263-1486, 1263-1520, 1263-1738, 1263-1764, 1264-1515, 1264-1700, 1268-1527, 1274-1728, 1285-1540, 1289-1729, 1294-1540, 1294-1542, 1295-1571, 1295-1729, 1296-1761, 1301-1729, 1305-1765, 1311-1713, 1312-1759, 1319-1729, 1325-1559, 1325-1729, 1327-1729, 1329-1729, 1332-1644, 1333-1729, 1334-1646, 1335-1558, 1335-1590, 1335-1729, 1341-1729, 1343-1626, 1347-1577, 1353-1592, 1359-1729, 1360-1542, 1361-1609, 1361-1699, 1362-1765, 1367-1542, 1374-1650, 1374-1729, 1380-1729, 1382-1729, 1384-1729, 1385-1664, 1385-1729, 1387-1592, 1387-1635, 1387-1729, 1387-1761, 1388-1729, 1393-1537, 1394-1740, 1395-1729, 1396-1729, 1399-1729, 1400-1729, 1401-
	1761, 1401-1765, 1403-1729, 1405-1684, 1407-1729, 1408-1729, 1409-1765, 1414-1765, 1419-1680, 1424-1671, 1426-1729, 1436-1714, 1442-1729, 1446-1659, 1447-1729, 1459-1765, 1461-1729, 1462-1729, 1463-1729, 1472-1755, 1474-1691, 1477-1727, 1478-1654, 1494-1690, 1497-1729, 1502-1711, 1502-1765, 1504-1765, 1505-1765, 1507-1729, 1523-1745, 1528-1729, 1529-1729, 1536-1729, 1537-1760, 1537-1765, 1543-1765, 1544-1765, 1547-1729, 1569-1729, 1574-1729, 1575-1729, 1581-1765, 1585-1764, 1588-1757, 1612-1765, 1623-1759, 1628-1765, 1634-1759, 1636-1765, 1642-1765, 1653-1729
63/ 7500687CB1/ 1695	1-47, 1-1695, 95-352, 128-393, 128-586, 128-601, 128-604, 128-611, 128-626, 131-484, 132-650, 133-653, 138-424, 140-625, 143-834, 151-348, 151-637, 155-636, 155-741, 157-586, 164-352, 167-442, 173-811, 174-427, 174-786, 174-823, 197-471, 197-922, 211-757, 211-773, 216-681, 217-730, 239-730, 252-351, 259-787, 263-814, 263-853, 265-717, 268-878, 269-506, 272-560, 273-419, 286-482, 294-800, 306-465, 307-926, 320-883, 325-883, 362-904, 364-872, 371-866, 372-643, 383-865, 389-1040, 396-871, 396-1078, 397-674, 398-635, 400-978, 406-957, 410-867, 410-919, 410-975, 412-891, 421-895, 427-904, 458-921, 458-1107, 461-734, 466-694, 466-1071, 469-1012, 471-905, 471-907, 472-734, 473-1149, 492-650, 494-910, 494-1092, 507-655, 519-793, 522-1212, 543-951, 543-1072, 543-1213, 548-708, 548-1152, 548-1228, 561-1412, 565-1101, 565-1106, 576-939, 577-1121, 583-1251, 584-862, 584-1067, 585-932, 592-862, 592-882, 592-1016, 598-1192, 599-1185, 606-1246, 611-833, 615-1215, 630-878, 637-794, 641-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1002, 653-1446, 655-928, 664-928, 669-1260, 675-1305, 683-1169, 687-1110, 692-1297, 692-1322, 694-973, 696-1110, 696-1171, 700-1199, 705-929, 707-954, 710-952, 710-1234, 711-952, 712-1034, 712-1036, 717-983, 725-939, 725-1389, 733-1504, 738-1224, 739-985, 739-1046, 746-1195, 750-1021, 762-1237, 780-1308, 810-1466, 822-1114, 825-1009, 829-1322, 839-1349, 844-1635, 863-1332, 882-1483, 893-1354, 899-1484, 910-1151, 910-1412, 914-1156, 914-1177, 918-1227, 937-1582, 954-1324, 955-1218, 960-1489, 984-1408, 1002-1483, 1009-1552, 1011-1333, 1011-1371, 1015-1185, 1018-1462, 1025-1354, 1045-1667, 1045-1669, 1051-1656, 1053-1315, 1066-1548, 1090-1674, 1105-1298, 1110-1695, 1117-1371, 1120-1405, 1124-1677, 1124-1695, 1129-1659, 1147-1674, 1153-1671, 1158-1432, 1167-1613, 1173-1656, 1178-1460, 1179-1671, 1180-1508, 1181-1690, 1187-1692, 1188-1319, 1189-1675, 1195-1671, 1197-1692, 1198-1674, 1214-1673, 1218-1565, 1218-1673, 1221-1674, 1225-1674, 1226-1693,
	1231-1684, 1232-1476, 1233-1673, 1235-1678, 1242-1673, 1242-1677, 1245-1647, 1249-1548, 1251-1695, 1259-1674, 1264-1485, 1266-1436, 1299-1560, 1308-1443, 1317-1677, 1327-1603, 1329-1660, 1332-1682, 1340-1592, 1349-1673, 1351-1674, 1427-1658, 1428-1668, 1428-1690, 1474-1673, 1573-1695, 1600-1674
64/ 7500688CB1/ 2202	1-47, 1-538, 1-2202, 127-405, 127-536, 131-578, 134-622, 136-410, 138-484, 188-272, 191-517, 199-607, 208-709, 238-489, 256-425, 258-360, 266-709, 269-387, 297-855, 305-804, 351-852, 359-582, 409-924, 414-904, 415-912, 442-637, 447-1098, 458-987, 478-925, 506-1139, 512-1108, 514-1037, 524-981, 525-1029, 531-979, 625-1157, 635-1093, 635-1108, 635-1111, 635-1118, 635-1133, 638-991, 640-1160, 645-931, 647-1132, 650-1341, 658-855, 658-1144, 662-1143, 662-1248, 664-1093, 671-859, 674-949, 680-1318, 681-934, 681-1293, 681-1330, 704-978, 704-1429, 718-1264, 718-1280, 723-1188, 724-1237, 746-1237, 759-858, 766-1294, 770-1321, 770-1360, 772-1224, 775-1385, 776-1013, 779-1067, 780-926, 793-989, 801-1307, 813-972, 814-1433, 827-1390, 832-1390, 869-1411, 871-1379, 878-1373, 879-1150, 890-1372, 891-1093, 896-1547, 903-1378, 903-1585, 904-1181, 905-1142, 907-1485, 913-1464, 917-1374, 917-1426, 917-1482, 919-1398, 928-1402, 934-1411, 965-1428, 965-1614, 968-1241, 973-1201, 973-1578,
	976-1519, 978-1412, 978-1414, 979-1241, 980-1656, 999-1157, 1001-1417, 1001-1599, 1014-1162, 1026-1300, 1029-1719, 1050-1458, 1050-1579, 1050-1720, 1055-1215, 1055-1659, 1055-1735, 1068-1919, 1072-1608, 1072-1613, 1083-1446, 1084-1628, 1090-1758, 1091-1369, 1091-1574, 1092-1439, 1099-1369, 1099-1389, 1099-1523, 1105-1699, 1106-1692, 1113-1753, 1118-1340, 1122-1722, 1137-1385, 1144-1301, 1148-1509, 1160-1953, 1162-1435, 1171-1435, 1176-1767, 1182-1812, 1190-1676, 1194-1617, 1199-1804, 1199-1829, 1201-1480, 1203-1617, 1203-1678, 1207-1706, 1212-1436, 1214-1461, 1217-1459, 1217-1741, 1218-1459, 1219-1541, 1219-1543, 1224-1490, 1232-1446, 1232-1896, 1240-2011, 1245-1731, 1246-1492, 1246-1553, 1253-1702, 1257-1528, 1269-1744, 1287-1815, 1317-1973, 1329-1621, 1332-1516, 1336-1829, 1346-1856, 1351-2142, 1370-1839, 1389-1990, 1400-1861, 1406-1991, 1417-1658, 1417-1919, 1421-1663, 1421-1684, 1425-1734, 1444-2089, 1461-1831, 1462-1725,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1467-1996, 1491-1915, 1509-1990, 1516-2059, 1518-1840, 1518-1878, 1522-1692, 1525-1969, 1532-1861, 1552-2174, 1552-2176, 1558-2163, 1560-1822, 1573-2055, 1597-2181, 1612-1805, 1617-2202, 1624-1878, 1627-1912, 1631-2184, 1631-2202, 1636-2166, 1654-2181, 1660-2178, 1665-1939, 1674-2120, 1680-2163, 1685-1967, 1686-2178, 1687-2015, 1688-2197, 1694-2199, 1695-1826, 1696-2182, 1702-2178, 1704-2199, 1705-2181, 1721-2180, 1725-2072, 1725-2180, 1728-2181, 1732-2181, 1733-2200, 1738-2191, 1739-1983, 1740-2180, 1742-2185, 1749-2180, 1749-2184, 1752-2154, 1756-2055, 1758-2202, 1766-2181, 1771-1992, 1773-1943, 1806-2067, 1815-1950, 1824-2184, 1834-2110, 1836-2167, 1839-2189, 1847-2099, 1856-2180, 1858-2181, 1934-2165, 1935-2175, 1935-2197, 1981-2180, 2080-2202, 2107-2181
65/ 7500697CB1/ 779	1-779, 43-225, 57-216, 59-189, 59-218, 60-218, 60-219, 62-218, 63-218, 66-218, 68-327, 80-348, 80-349, 82-218, 218-472, 218-738, 218-756, 218-774, 218-778, 219-731, 220-483, 221-516, 222-504, 223-411, 226-447, 243-488, 245-499, 246-748, 254-539, 267-541, 274-779, 284-549, 287-546, 298-779, 299-764, 301-484, 302-577, 314-763, 315-764, 324-763, 326-561, 337-601, 344-767, 359-554, 360-779, 365-758, 374-768, 387-641, 388-779, 402-764, 405-768, 412-767, 421-763, 433-747, 435-671, 435-695, 435-748, 440-770, 445-760, 447-779, 457-768, 469-762, 473-771, 474-770, 476-761, 481-777, 504-768, 530-779, 534-777, 576-763, 588-779, 606-774, 651-767, 662-779, 668-764
66/ 7500709CB1/ 3104	1-501, 1-3052, 106-366, 278-708, 280-429, 280-494, 280-864, 283-566, 283-1042, 283-1044, 284-1014, 291-921, 292-546, 302-552, 306-506, 306-547, 306-591, 308-920, 309-783, 309-935, 309-1118, 310-601, 310-1165, 313-660, 315-772, 317-890, 321-969, 322-945, 323-1005, 394-636, 458-1213, 464-976, 532-1230, 534-1144, 535-1049, 588-1132, 602-1113, 622-1259, 625-1256, 643-947, 655-1243, 657-1280, 657-1286, 658-1208, 678-1050, 691-1325, 692-1262, 715-966, 715-1583, 732-1277, 735-897, 737-1006, 741-1509, 747-1463, 757-1323, 757-1325, 765-1401, 783-1405, 793-1382, 804-1413, 811-1349, 814-1531, 820-1286, 827-1504, 845-1065, 845-1390, 845-1420, 845-1455, 845-1636, 849-1388, 849-1476, 860-1171, 860-1361, 881-1535, 881-1574, 881-1610, 898-1581, 900-1072, 904-1456, 927-1446, 932-1609, 943-1244, 948-1203, 954-1253, 954-1310, 959-1635, 962-1605, 966-1583, 967-1567, 984-1580, 990-1580, 991-1600, 998-1284, 1004-1648, 1006-1322, 1009-1501, 1024-1152, 1024-1255, 1050-1636, 1057-1542, 1066-1645,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1085-1304, 1088-1565, 1124-1602, 1160-1531, 1175-1390, 1181-1838, 1181-1841, 1181-1864, 1181-1869, 1250-2098, 1255-1554, 1355-1634, 1374-1624, 1656-1925, 1656-2254, 1661-1903, 1661-1956, 1661-2243, 1661-2438, 1663-2343, 1678-2188, 1694-1776, 1698-1946, 1698-2046, 1699-1904, 1699-1950, 1707-1963, 1711-2123, 1712-1895, 1712-1929, 1724-2255, 1733-2464, 1748-2110, 1759-1934, 1761-2059, 1764-2158, 1766-2128, 1767-2250, 1770-2347, 1790-2459, 1791-2369, 1795-2097, 1796-2097, 1809-2078, 1812-2087, 1822-2096, 1822-2100, 1828-2158, 1850-2100, 1863-2603, 1886-2570, 1888-2546, 1891-2493, 1900-2141, 1901-2524, 1902-2141, 1912-2385, 1917-2186, 1923-2168, 1924-2525, 1928-2217, 1929-2207, 1937-2144, 1937-2420, 1959-2223, 1965-2225, 1965-2458, 1966-2215, 1966-2466, 1970-2679, 1975-2665, 1988-2589, 1988-2596, 1990-2250, 1996-2308, 1996-2309, 2001-2550, 2002-2401, 2013-2517, 2014-2262, 2018-2233, 2018-2511, 2019-2138, 2019-2295, 2020-2493,
	2021-2303, 2024-2318, 2026-2224, 2026-2273, 2027-2243, 2027-2607, 2052-2214, 2057-2674, 2062-2273, 2075-2318, 2078-2306, 2082-2982, 2083-2293, 2083-2827, 2090-2712, 2101-2726, 2106-2414, 2108-2381, 2123-2363, 2124-2354, 2124-2366, 2124-2394, 2127-2773, 2128-2654, 2128-2674, 2128-2682, 2129-2674, 2151-2393, 2153-2389, 2164-2458, 2177-2411, 2199-2554, 2207-2856, 2212-2522, 2225-3070, 2226-2417, 2229-2460, 2258-2815, 2260-2493, 2268-2909, 2271-2893, 2273-2460, 2284-2470, 2287-2961, 2307-2603, 2321-2901, 2326-2576, 2326-2586, 2347-2563, 2352-2991, 2355-3024, 2358-3014, 2367-2614, 2367-2955, 2367-2978, 2367-3037, 2375-2623, 2375-2653, 2397-3041, 2398-2684, 2418-2828, 2418-2852, 2419-2671, 2425-2989, 2433-2736, 2442-3014, 2443-3030, 2455-2753, 2456-2740, 2457-2654, 2470-2758, 2470-2990, 2481-3037, 2483-2913, 2488-2917, 2492-2865, 2498-2736, 2503-3056, 2509-3010, 2519-2692, 2527-3028, 2541-3037, 2554-2836, 2557-2790, 2561-3032,
	2574-2817, 2574-2825, 2579-3034, 2583-2915, 2587-2817, 2588-2804, 2592-3037, 2595-2830, 2595-3059, 2599-3035, 2601-3060, 2602-3019, 2605-3037, 2624-3037, 2629-2932, 2631-3037, 2636-3032, 2640-2910, 2645-3056, 2649-3104, 2652-3080, 2661-2912, 2664-3059, 2686-3061, 2688-2957, 2691-3062, 2697-2929, 2701-3056, 2706-3063, 2707-3060, 2708-3030, 2713-3060, 2717-3055, 2721-3056, 2731-3060, 2732-3032, 2732-3036, 2735-3061, 2745-3030, 2747-3002, 2767-2991, 2788-3030, 2793-3063, 2796-3064, 2799-3032, 2801-3034, 2824-3032, 2827-3037, 2838-3060, 2838-3088, 2846-3060, 2861-3031, 2898-3032

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
67/ 7500711CB1/ 3233	1-501, 1-3181, 106-366, 282-728, 288-812, 288-817, 299-812, 299-817, 302-552, 302-619, 307-817, 307-978, 308-503, 309-522, 309-564, 309-595, 309-783, 309-821, 310-601, 311-563, 315-772, 352-814, 394-636, 417-685, 417-810, 425-810, 466-812, 510-773, 699-1178, 728-788, 854-1508, 854-1547, 871-1554, 877-1429, 879-1745, 900-1419, 927-1226, 927-1283, 932-1608, 939-1556, 940-1540, 943-1718, 950-1744, 957-1553, 959-1633, 963-1553, 964-1573, 969-1696, 971-1257, 977-1627, 978-1695, 979-1295, 996-1599, 997-1125, 1005-1697, 1018-1635, 1024-1650, 1030-1515, 1039-1618, 1039-1701, 1045-1658, 1051-1197, 1056-1749, 1061-1538, 1061-1721, 1073-1774, 1092-1747, 1096-1891, 1097-1575, 1111-1736, 1119-1929, 1133-1504, 1145-1781, 1148-1363, 1150-1931, 1183-1871, 1194-1794, 1195-1853, 1207-1693, 1208-1610, 1209-1610, 1210-1893, 1213-1481, 1214-1811, 1218-1514, 1222-1809, 1228-1527, 1232-1700, 1232-1766, 1234-1457, 1235-1791, 1235-1907, 1244-1575, 1253-1837, 1258-1893, 1262-1543,
	1262-1789, 1269-1892, 1272-1796, 1273-1818, 1273-1911, 1277-1811, 1281-1701, 1290-1909, 1291-1628, 1295-2024, 1303-1832, 1308-2117, 1311-1985, 1312-2015, 1313-2210, 1323-1545, 1323-1932, 1332-1621, 1347-1597, 1348-1970, 1364-1941, 1369-1631, 1392-1974, 1397-2090, 1398-1986, 1422-2053, 1423-1663, 1434-1882, 1444-1728, 1444-2127, 1446-2301, 1457-1697, 1457-1726, 1460-1996, 1462-1943, 1463-2153, 1464-1948, 1464-2122, 1464-2152, 1467-1936, 1475-1639, 1479-2041, 1479-2183, 1480-2058, 1485-1799, 1490-2139, 1491-2043, 1503-2066, 1505-1783, 1522-2237, 1553-2213, 1557-1803, 1557-1813, 1564-1806, 1571-2380, 1573-2339, 1598-2077, 1602-2213, 1608-1927, 1608-2300, 1609-2188, 1611-2183, 1612-2125, 1624-1799, 1626-2371, 1628-2198, 1628-2318, 1635-2181, 1654-1923, 1654-2313, 1657-2475, 1670-2015, 1678-2411, 1681-1980, 1703-2341, 1713-2239, 1715-1996, 1747-1996, 1748-2257, 1748-2327, 1765-1994, 1770-2173, 1770-2227, 1775-2135, 1776-2054,
	1785-2054, 1785-2383, 1790-2032, 1790-2085, 1790-2372, 1790-2567, 1792-2472, 1806-1946, 1807-2317, 1810-2001, 1823-1905, 1827-2075, 1827-2175, 1828-2033, 1836-2092, 1840-2252, 1853-2384, 1862-2593, 1873-2392, 1879-2089, 1888-2063, 1890-2188, 1893-2287, 1895-2257, 1896-2379, 1899-2476, 1919-2595, 1920-2498, 1924-2226, 1925-2226, 1938-2207, 1951-2225, 1951-2229, 1957-2287, 1979-2229, 1992-2732, 2015-2699, 2017-2675, 2020-2624, 2029-2270, 2030-2656, 2031-2270, 2041-2514, 2046-2315, 2052-2297, 2053-2654, 2057-2346, 2058-2297, 2058-2336, 2066-2273, 2066-2549, 2088-2352, 2094-2354, 2094-2587, 2095-2344, 2095-2595, 2099-2808, 2104-2794, 2104-2841, 2105-2644, 2117-2718, 2117-2725, 2119-2379, 2125-2437, 2125-2438, 2126-2437, 2130-2679, 2131-2530, 2142-2646, 2143-2391, 2147-2362, 2147-2640, 2148-2267, 2148-2424, 2149-2622, 2150-2432, 2153-2447, 2155-2353, 2155-2402, 2156-2372, 2156-2736, 2165-2851, 2181-2343, 2183-2860, 2186-2803,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	2191-2402, 2198-2881, 2202-2816, 2204-2447, 2207-2435, 2211-3111, 2212-2422, 2212-2956, 2219-2841, 2230-2855, 2232-2482, 2235-2543, 2237-2510, 2238-2810, 2247-2925, 2249-2949, 2252-2492, 2253-2483, 2253-2495, 2253-2523, 2256-2902, 2257-2783, 2257-2803, 2257-2811, 2258-2803, 2279-2960, 2280-2522, 2282-2518, 2293-2587, 2299-2909, 2306-2540, 2314-2852, 2324-2650, 2328-2683, 2332-2785, 2332-2968, 2336-2985, 2341-2653, 2346-2936, 2354-3199, 2355-2546, 2357-2808, 2358-2591, 2358-2624, 2386-2574, 2387-2605, 2387-2944, 2388-3109, 2389-2624, 2397-3038, 2400-3022, 2402-2589, 2407-2677, 2409-2677, 2413-2603, 2416-3090, 2436-2732, 2450-3030, 2455-2705, 2455-2715, 2476-2692, 2479-3094, 2481-2747, 2481-3120, 2484-3153, 2487-3143, 2496-2743, 2496-3084, 2496-3107, 2496-3166, 2504-2752, 2504-2782, 2526-3170, 2527-2813, 2547-2957, 2547-2981, 2548-2800, 2554-3118, 2562-2865, 2571-3143, 2572-3159, 2572-3166, 2584-2882, 2585-2869, 2585-2872,
	2585-2888, 2586-2783, 2599-2887, 2599-3119, 2610-3166, 2612-3042, 2617-3046, 2619-2877, 2621-2994, 2627-2865, 2632-3185, 2638-3139, 2641-3166, 2647-2929, 2648-2821, 2656-3157, 2670-3166, 2677-3166, 2683-2965, 2686-2919, 2690-3161, 2703-2946, 2703-2954, 2708-3163, 2712-3044, 2716-2946, 2717-2933, 2721-3166, 2722-3161, 2724-2959, 2724-3188, 2728-3164, 2730-3189, 2731-3148, 2734-3166, 2746-3166, 2753-3166, 2758-3061, 2760-3166, 2765-3161, 2769-3039, 2773-3161, 2774-3185, 2777-3161, 2778-3233, 2781-3209, 2790-3041, 2793-3188, 2796-3161, 2796-3164, 2802-3184, 2809-3055, 2815-3190, 2816-3161, 2817-3086, 2817-3161, 2820-3191, 2823-3184, 2826-3058, 2828-3164, 2830-3185, 2835-3192, 2836-3189, 2837-3159, 2839-3161, 2842-3189, 2846-3184, 2850-3185, 2853-3166, 2860-3006, 2860-3189, 2861-3161, 2861-3165, 2864-3190, 2871-3161, 2874-3159, 2874-3163, 2876-3131, 2877-3166, 2879-3161, 2896-3120, 2896-3166, 2917-3159, 2920-3161, 2922-3192,
	2924-3161, 2925-3193, 2927-3166, 2928-3161, 2929-3190, 2930-3163, 2943-3161, 2953-3161, 2956-3166, 2967-3189, 2967-3217, 2975-3189, 2990-3160, 3027-3161, 3042-3184, 3043-3184, 3077-3166
68/ 7500723CB1/ 1808	1-198, 1-1808, 47-742, 57-338, 64-149, 72-200, 208-417, 208-441, 208-485, 208-723, 208-733, 208-888, 219-927, 219-936, 234-792, 234-902, 243-431, 248-507, 253-646, 261-793, 275-554, 280-823, 290-925, 292-560, 303-613, 308-702, 310-526, 317-1050, 321-498, 322-586, 323-642, 323-904, 327-945, 331-977, 333-951, 335-993, 336-1037, 338-1005, 343-874, 347-600, 347-996, 348-460, 376-706, 380-628, 386-549, 401-951, 407-892, 420-818, 421-800, 433-677, 435-688, 444-732, 444-737, 453-758, 463-1100, 466-943, 467-905, 482-762, 482-769, 484-818, 487-722, 494-704, 497-700, 508-1047, 508-1278, 513-1263, 514-701, 515-1155, 521-914, 530-1264, 531-777, 534-790, 538-1192, 539-782, 541-1090, 543-647, 543-1041, 548-1080, 549-1076, 553-1237, 556-821, 556-1100, 559-1032, 570-747, 572-862, 574-1004, 574-1012, 577-1087, 581-855, 585-1151, 586-825, 592-1108, 596-854, 597-1201, 599-1071, 601-875, 601-1177, 601-1257, 605-1240, 616-774, 618-1314, 626-876, 626-898, 628-875, 629-1220, 629-1228, 635-996, 637-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	900, 637-1311, 638-838, 638-839, 642-1001, 645-864, 645-913, 645-914, 646-919, 652-876, 653-879, 655-923, 658-1264, 662-915, 665-889, 674-1018, 678-889, 682-969, 682-1291, 685-1278, 687-1312, 699-930, 699-935, 699-985, 700-1209, 715-844, 716-1231, 716-1311, 716-1315, 723-979, 728-969, 736-1290, 738-968, 739-1351, 747-1450, 747-1589, 763-971, 771-1041, 771-1231, 772-1026, 775-1152, 775-1261, 776-1044, 784-1085, 788-1021, 796-1614, 801-1073, 802-1096, 803-933, 803-1087, 804-1099, 806-971, 807-1535, 807-1551, 814-1416, 830-1078, 830-1137, 830-1457, 831-1397, 833-1151, 834-1069, 841-1121, 841-1155, 843-1026, 848-1422, 854-1121, 858-1112, 858-1118, 860-1424, 861-1066, 861-1081, 876-1119, 879-1151, 879-1163, 880-1246, 881-1057, 882-1455, 888-1148, 893-1452, 894-1124, 894-1295, 897-1172, 898-1160, 898-1200, 899-1509, 900-1164, 903-1298, 912-1489, 916-1173, 918-1156, 919-1160, 919-1161, 919-1190, 921-1113, 922-1189, 929-1224, 931-1726, 935-1471, 936-1588, 950-1196, 951-1503, 962-1794,
	963-1181, 967-1252, 969-1391, 970-1474, 976-1259, 978-1228, 979-1733, 982-1068, 983-1271, 986-1650, 988-1286, 988-1304, 990-1794, 992-1549, 995-1573, 997-1258, 997-1280, 997-1364, 997-1752, 999-1230, 999-1489, 999-1651, 1001-1529, 1006-1734, 1007-1274, 1009-1295, 1010-1616, 1012-1761, 1015-1275, 1018-1292, 1024-1599, 1025-1283, 1028-1236, 1028-1313, 1028-1447, 1029-1299, 1030-1261, 1033-1315, 1033-1349, 1033-1585, 1043-1501, 1045-1305, 1045-1607, 1049-1265, 1049-1723, 1054-1349, 1057-1279, 1057-1280, 1057-1311, 1057-1609, 1058-1302, 1065-1303, 1065-1397, 1068-1316, 1068-1580, 1070-1289, 1071-1740, 1072-1746, 1075-1312, 1076-1346, 1081-1749, 1083-1333, 1090-1485, 1093-1644, 1100-1585, 1102-1573, 1106-1653, 1106-1756, 1107-1359, 1111-1410, 1112-1350, 1113-1801, 1115-1384, 1116-1407, 1117-1384, 1120-1689, 1120-1781, 1124-1369, 1124-1400, 1126-1440, 1127-1395, 1127-1478, 1129-1298, 1129-1363, 1129-1392, 1129-1585, 1131-1406, 1133-1787, 1136-
	1621, 1143-1451, 1144-1787, 1145-1386, 1147-1801, 1149-1704, 1152-1689, 1153-1366, 1154-1338, 1154-1344, 1154-1375, 1154-1404, 1154-1406, 1154-1430, 1154-1446, 1154-1454, 1154-1801, 1163-1427, 1165-1284, 1165-1394, 1166-1408, 1170-1705, 1170-1799, 1173-1480, 1177-1444, 1179-1373, 1182-1801, 1184-1727, 1198-1764, 1199-1434, 1201-1791, 1222-1484, 1226-1628, 1230-1778, 1233-1415, 1233-1505, 1233-1550, 1233-1783, 1239-1445, 1240-1752, 1242-1791, 1246-1786, 1251-1440, 1252-1638, 1257-1803, 1272-1548, 1278-1506, 1281-1801, 1294-1519, 1303-1808, 1305-1537, 1306-1624, 1316-1525, 1316-1791, 1319-1801, 1330-1587, 1333-1587, 1338-1808, 1348-1805, 1366-1611, 1375-1553, 1375-1670, 1387-1787, 1393-1801, 1394-1650, 1396-1801, 1401-1801, 1404-1646, 1414-1801, 1416-1651, 1417-1755, 1419-1801, 1420-1668, 1420-1680, 1420-1808, 1423-1668, 1423-1670, 1423-1674, 1430-1674, 1436-1684, 1437-1808, 1438-1808, 1441-1716, 1454-1788, 1456-1729, 1458-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1720, 1458-1788, 1466-1755, 1484-1758, 1484-1776, 1484-1798, 1486-1767, 1487-1796, 1489-1801, 1493-1733, 1493-1747, 1505-1740, 1505-1751, 1505-1761, 1505-1762, 1505-1801, 1531-1711, 1531-1746, 1535-1780, 1540-1790, 1540-1801, 1545-1801, 1548-1798, 1555-1801, 1579-1791, 1579-1801, 1597-1787, 1597-1796, 1597-1801, 1600-1801, 1606-1756, 1608-1808, 1624-1791, 1629-1801, 1634-1791, 1651-1791, 1663-1801, 1665-1801, 1667-1788, 1667-1798, 1668-1802, 1669-1799, 1670-1803, 1672-1787, 1672-1801, 1674-1801, 1676-1787, 1699-1798, 1699-1801, 1700-1791, 1705-1799, 1706-1798, 1710-1801, 1747-1801
69/ 7500764CB1/ 1672	1-1672, 4-1672, 99-702, 157-624, 157-657, 157-695, 157-740, 158-815, 197-749, 205-715, 230-906, 234-524, 254-824, 265-686, 269-897, 270-686, 293-911, 298-822, 308-849, 313-875, 316-857, 337-883, 337-886, 338-816, 348-958, 369-510, 370-873, 410-681, 415-811, 417-1000, 417-1051, 426-1016, 427-642, 427-670, 427-949, 431-994, 431-1111, 439-949, 442-949, 454-1111, 470-760, 473-870, 473-1062, 476-1044, 476-1099, 499-586, 503-949, 523-916, 524-771, 524-884, 526-1014, 536-644, 541-767, 551-1178, 559-1032, 564-1193, 594-903, 597-1077, 603-726, 604-1157, 605-1191, 611-1300, 623-1203, 630-1196, 634-1334, 636-1206, 640-1236, 645-1204, 659-1239, 660-1276, 663-1330, 666-779, 666-1260, 668-1359, 675-875, 682-1326, 697-855, 701-1088, 702-1294, 704-1278, 716-1161, 735-885, 735-1026, 737-948, 737-955, 748-1044, 772-876, 784-1394, 795-1032, 797-1431, 802-1149, 816-1543, 821-1034, 834-1132, 839-1106, 842-1526, 848-1466, 853-1455, 859-1133, 867-1128, 872-1483, 876-1165, 882-1472, 883-1214,
	888-1306, 893-1580, 897-1143, 902-1359, 912-1558, 913-1509, 914-1116, 914-1345, 914-1514, 917-1203, 917-1487, 919-1518, 922-1542, 939-1526, 950-1545, 971-1230, 971-1671, 974-1607, 977-1508, 979-1493, 983-1483, 989-1526, 991-1455, 998-1097, 1005-1559, 1014-1558, 1014-1649, 1015-1582, 1018-1582, 1022-1269, 1025-1578, 1031-1549, 1040-1494, 1058-1235, 1072-1325, 1072-1512, 1083-1672, 1096-1456, 1139-1672, 1146-1672, 1149-1453, 1152-1411, 1157-1402, 1159-1639, 1172-1410, 1240-1478, 1245-1446, 1255-1575
70/ 7500772CB1/ 1347	1-1338, 1-1347, 3-139, 16-131, 42-315, 123-416, 156-334, 157-521, 157-628, 161-419, 169-734, 170-422, 185-456, 185-731, 190-673, 216-435, 227-676, 238-874, 248-418, 248-502, 248-750, 251-814, 254-845, 286-696, 295-843, 299-839, 299-861, 305-613, 305-941, 318-738, 318-953, 325-677, 334-831, 339-470, 342-601, 343-616, 345-1015, 363-927, 364-963, 365-614, 381-839, 381-1004, 384-1010, 390-916, 396-679, 400-977, 412-712, 431-762, 434-719, 435-720, 443-693, 443-739, 445-901, 457-1012, 466-746, 472-772, 479-732, 479-1024, 482-741, 483-809, 484-727, 494-1003, 495-962, 496-783, 502-1061, 504-690, 509-1127, 515-631, 516-885, 520-1158, 520-1320, 522-777, 522-952, 527-814, 536-792, 536-1079, 536-1104, 541-754, 543-741, 543-771, 548-689, 550-824, 554-768, 554-773, 555-695, 555-1082, 556-844, 562-1105, 563-1113, 564-804, 565-807, 568-815, 568-844, 570-1078, 570-1131, 577-719, 581-862, 581-900, 586-632, 586-836, 588-865, 596-905, 596-1106, 596-1206, 598-806, 598-813, 598-842, 598-862,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	598-876, 599-1131, 602-840, 602-895, 602-1172, 605-805, 607-1107, 617-837, 620-884, 621-865, 622-1189, 623-870, 623-873, 623-926, 623-1134, 627-1067, 629-1082, 635-881, 637-1198, 638-850, 640-877, 664-925, 665-926, 669-912, 669-919, 669-920, 669-927, 669-1238, 670-867, 670-911, 673-946, 678-898, 678-1035, 679-956, 689-931, 691-989, 699-963, 700-927, 700-988, 701-996, 706-925, 714-835, 714-1129, 724-1043, 724-1157, 727-948, 727-1132, 727-1159, 728-994, 730-937, 730-1038, 733-1194, 734-821, 734-1015, 734-1347, 747-1347, 748-840, 752-987, 759-1347, 761-1182, 767-1040, 771-1027, 778-1347, 785-1347, 793-1094, 794-1063, 804-1006, 807-1100, 824-1078, 825-1002, 843-1063, 844-1012, 847-1068, 847-1084, 854-1012, 919-1153, 970-1137
71/ 7501350CB1/ 729	1-693, 120-683, 153-412, 206-682, 259-519, 260-496, 260-509, 260-683, 260-687, 261-682, 262-687, 269-571, 269-690, 271-529, 271-561, 273-418, 276-490, 276-700, 277-492, 281-635, 282-729, 284-682, 293-475, 296-587, 301-626, 303-573, 306-682, 307-540, 307-544, 307-682, 316-572, 321-616, 322-602, 324-647, 327-682, 328-572, 329-686, 330-515, 338-578, 338-687, 350-638, 362-561, 363-665, 371-650, 372-682, 378-635, 386-682, 406-685, 409-675, 431-682, 460-682, 475-681, 508-698
72/ 7506396CB1/ 1188	1-305, 1-1182, 219-420, 219-432, 219-472, 219-547, 264-1118, 287-657, 318-498, 324-657, 331-1126, 361-1123, 375-579, 375-657, 387-634, 387-641, 397-1123, 418-628, 418-646, 426-657, 443-657, 454-680, 459-657, 657-866, 657-889, 657-1135, 657-1162, 657-1188, 658-945, 659-898, 660-1188, 662-1188, 676-980, 705-1183, 708-964, 708-1144, 708-1164, 712-1171, 713-1188, 717-1174, 720-1169, 722-1006, 722-1010, 722-1013, 722-1186, 724-1161, 725-1168, 728-915, 736-1165, 748-1170, 753-1187, 758-1165, 759-1041, 759-1102, 763-1171, 764-1167, 767-1173, 768-1173, 769-1168, 769-1171, 769-1188, 776-1168, 779-1171, 799-1044, 801-1166, 802-1065, 803-1169, 804-1167, 812-1081, 813-1166, 829-944, 830-1186, 832-1165, 832-1171, 856-1171, 885-1154, 885-1182, 908-1173, 940-1167, 942-1171, 954-1171, 961-1166, 985-1123, 1047-1168
73/ 7505917CB1/ 1886	1-334, 1-501, 1-1879, 32-327, 99-284, 266-1000, 279-977, 281-988, 281-1114, 297-530, 311-950, 320-572, 320-637, 320-742, 320-771, 320-809, 326-908, 329-507, 329-735, 329-790, 329-825, 329-848, 329-850, 329-881, 329-938, 338-801, 349-810, 351-773, 351-781, 352-758, 355-742, 363-772, 371-823, 478-713, 480-756, 520-820, 580-979, 586-812, 586-1139, 614-849, 655-912, 673-934, 674-927, 682-1245, 707-943, 710-984, 714-924, 714-1134, 755-992, 808-1051, 809-1205, 836-1080, 852-1388, 854-1148, 878-1099, 878-1110, 884-1128, 884-1265, 887-1281, 894-1174, 896-1140, 904-1198, 910-1129, 943-1178, 962-1223, 965-1225, 968-1219, 981-1229, 989-1249, 998-1246, 1019-1318, 1027-1304, 1033-1339, 1052-1307, 1056-1305, 1075-1347, 1088-1353, 1111-1351, 1112-1233, 1120-1301, 1129-1378, 1165-1388, 1175-1297, 1218-1388, 1246-1388, 1311-1574, 1387-1613, 1387-1627, 1387-1635, 1387-1640, 1387-1827, 1387-1849, 1387-1873, 1387-1885, 1390-1866, 1395-1649, 1399-1691, 1406-1654, 1410-1865,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1411-1673, 1418-1886, 1436-1883, 1436-1886, 1443-1682, 1447-1885, 1448-1868, 1450-1866, 1461-1747, 1463-1885, 1472-1870, 1479-1800, 1480-1747, 1481-1692, 1481-1827, 1482-1872, 1495-1870, 1510-1585, 1523-1740, 1523-1830, 1529-1767, 1529-1830, 1529-1850, 1552-1871, 1572-1867, 1580-1837, 1592-1815, 1592-1881, 1613-1869, 1620-1878, 1625-1876, 1638-1883, 1638-1884, 1639-1886, 1641-1886, 1646-1815, 1653-1867, 1662-1886, 1683-1867, 1683-1886, 1686-1816, 1687-1886, 1704-1886, 1739-1886, 1744-1883
74/ 7500701CB1/ 1067	1-22, 1-265, 1-1067, 23-278, 23-380, 23-913, 25-138, 25-291, 25-307, 25-311, 26-257, 26-266, 26-276, 26-306, 27-296, 30-334, 31-295, 31-314, 32-292, 33-301, 34-274, 34-439, 35-227, 39-298, 40-271, 41-315, 42-262, 42-295, 44-324, 45-344, 47-329, 48-262, 57-306, 57-308, 62-353, 63-230, 63-336, 63-358, 63-382, 64-338, 65-312, 70-311, 72-334, 72-369, 76-346, 86-369, 88-380, 101-363, 108-359, 116-370, 117-418, 123-226, 336-479, 363-911, 381-911, 391-691, 450-894, 451-894, 466-866, 467-911, 472-893, 472-894, 504-913, 508-896, 509-866, 512-694, 519-893, 536-795, 634-860, 639-894, 660-901, 662-911, 666-894, 667-895, 815-894
75/ 7500702CB1/ 1220	1-22, 1-265, 1-1220, 23-278, 23-380, 23-1066, 25-138, 25-291, 25-307, 25-311, 26-257, 26-266, 26-276, 26-306, 27-296, 30-334, 31-295, 31-314, 32-292, 33-301, 34-274, 35-227, 39-298, 40-271, 41-315, 42-262, 42-295, 44-324, 45-344, 47-329, 48-262, 57-306, 57-308, 62-353, 63-230, 63-336, 63-358, 63-382, 64-338, 65-312, 70-311, 72-334, 72-369, 76-346, 86-358, 86-369, 88-380, 101-363, 116-370, 123-226, 293-556, 364-565, 379-1018, 382-678, 382-741, 382-910, 382-985, 382-1036, 382-1045, 385-950, 390-629, 392-665, 392-675, 395-660, 396-628, 407-635, 407-973, 441-681, 450-947, 450-952, 451-750, 452-733, 452-764, 453-978, 457-719, 457-726, 459-978, 461-982, 465-720, 470-1047, 471-737, 472-1047, 475-1064, 477-982, 511-656, 515-826, 518-774, 526-1064, 544-842, 544-1064, 555-815, 571-1046, 575-1047, 581-825, 592-792, 592-868, 593-1054, 595-1051, 598-888, 599-1064, 602-1041, 603-1047, 606-1046, 608-1043, 611-1064, 617-842, 617-1045, 618-992, 618-1064, 619-1045, 619-1046, 620-1046, 620-
	1064, 623-1017, 628-1044, 633-1046, 633-1064, 635-1047, 637-867, 640-1046, 641-1022, 646-995, 649-886, 650-1054, 651-1052, 665-847, 665-1017, 676-1048, 676-1050, 680-1064, 687-921, 689-946, 690-1047, 694-1046, 697-1051, 705-1045, 709-1047, 723-1038, 742-1051, 755-1034, 788-1013, 789-1051, 795-1046, 812-1050, 815-1064, 817-1046, 817-1054, 822-1038, 828-1050, 839-1064, 914-1044, 914-1045, 970-1045

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
76/ 6044343CB1/ 3280	1-99, 1-134, 1-263, 1-534, 1-700, 6-134, 7-134, 24-134, 262-1039, 264-1127, 399-1204, 457-1204, 462-1204, 482-1190, 683-1198, 768-1467, 789-1265, 813-1474, 832-901, 832-981, 832-982, 841-879, 851-879, 858-901, 890-978, 890-1040, 947-982, 956-1025, 964-998, 973-998, 1003-1653, 1046-1080, 1047-1085, 1433-2049, 1476-2048, 1520-2138, 1719-1975, 1719-1980, 1719-2231, 1730-2364, 1790-2034, 1963-2471, 2028-2331, 2030-2650, 2075-2340, 2134-2479, 2134-2795, 2162-2567, 2169-2389, 2275-2895, 2298-2530, 2320-2997, 2321-2580, 2359-2649, 2359-2778, 2359-2801, 2359-2818, 2359-2819, 2359-2865, 2359-2882, 2359-2910, 2359-2921, 2359-2936, 2359-2957, 2359-2967, 2359-2988, 2359-3033, 2359-3072, 2369-2622, 2370-2619, 2394-3044, 2398-2673, 2400-3106, 2453-3197, 2577-2827, 2581-3270, 2584-2827, 2584-2835, 2584-2836, 2584-2968, 2584-3032, 2613-3271, 2630-3280, 2639-3250, 2652-3102, 2654-3102, 2660-3280, 2664-3280, 2669-3276, 2679-3280, 2689-3280, 2691-3280,
	2720-3276, 2738-3085, 2738-3137, 2738-3263, 2747-3279, 2753-3091, 2754-3020, 2764-3280, 2771-3276, 2776-3280, 2779-3280, 2806-2957, 2833-3276, 2839-3276, 2840-3276, 2853-3276, 2863-3219, 2915-3276, 2924-3276, 2954-3276, 2969-3187, 3225-3274
77/ 7503990CB1/ 5567	1-5117, 1302-1556, 1443-1935, 1609-2295, 1978-2209, 1978-2638, 2069-2414, 2128-2590, 2153-2574, 2180-2624, 2253-2545, 2299-2404, 2299-2467, 2443-2729, 2466-3025, 2466-3039, 2556-2819, 3382-4042, 3402-3939, 3442-4041, 3454-4142, 3456-3961, 3466-3986, 3470-4086, 3475-4180, 3535-3788, 3536-3803, 3536-4071, 3536-4094, 3544-3788, 3546-3923, 3547-4063, 3547-4098, 3550-4202, 3555-4187, 3557-4060, 3558-4205, 3560-4099, 3592-3803, 3615-3910, 3619-4182, 3625-3817, 3632-4177, 3648-3933, 3676-4220, 3677-4132, 3681-4455, 3683-4227, 3690-4331, 3698-4278, 3698-4358, 3700-3946, 3719-4372, 3725-4111, 3728-4001, 3741-4122, 3746-4368, 3754-4419, 3771-4434, 3772-4268, 3773-4070, 3776-4022, 3777-4270, 3777-4369, 3781-4006, 3783-4070, 3800-4082, 3807-4242, 3813-4431, 3814-4085, 3828-4409, 3829-4368, 3859-4430, 3863-4490, 3871-4413, 3883-4432, 3889-4164, 3894-4468, 3895-4421, 3904-4185, 3914-4266, 3940-4378, 3968-4263, 3972-4251, 3983-4456, 3984-
	4127, 3985-4184, 3994-4070, 4016-4268, 4029-4616, 4043-4430, 4048-4608, 4053-4672, 4054-4669, 4087-4674, 4102-4264, 4102-4666, 4130-4381, 4140-4291, 4154-4428, 4154-4679, 4156-4622, 4171-4434, 4172-4424, 4201-4459, 4201-4700, 4202-4684, 4212-4501, 4231-4489, 4231-4723, 4234-4679, 4252-4498, 4279-4882, 4284-4412, 4285-4542, 4313-4496, 4314-4689, 4337-4462, 4386-4968, 4395-4663, 4418-4785, 4447-4746, 4450-4927, 4451-4900, 4453-4743, 4461-4895, 4466-4606, 4467-4731, 4487-4728, 4494-4625, 4494-4670, 4494-4750, 4525-4615, 4529-4791, 4570-4679, 4627-4858, 4627-5097, 4641-5117, 4669-4927, 4672-5095, 4681-4968, 4682-5145, 4684-5128, 4697-4846, 4742-4983, 4761-4993, 4767-5083, 4775-5043, 4789-5145, 4802-5083, 4813-4997, 4813-5083, 4813-5086, 4841-5104, 4856-5083, 4865-5145, 4867-5109, 4872-5145, 4882-5097, 4892-5083, 4896-5066, 4898-5083, 4920-5128, 4921-5083, 4924-5083, 4967-5083, 4976-5145, 5033-5567, 5260-5567

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
78/ 7504655CB1/ 1263	1-462, 1-1263, 491-591, 509-924, 516-1067, 521-976, 529-1052, 532-806, 543-679, 543-991, 551-927, 559-976, 574-1067, 605-864, 637-1056, 638-858, 638-883, 638-917, 638-972, 638-1018, 638-1043, 638-1049, 638-1056, 638-1063, 638-1064, 638-1065, 638-1068, 638-1069, 640-934, 641-1064, 641-1067, 642-928, 642-1072, 650-912, 651-1047, 652-919, 653-1049, 655-1051, 657-951, 659-1052, 660-937, 663-988, 664-1010, 665-1069, 666-834, 668-1067, 669-903, 674-1056, 677-965, 679-899, 679-1052, 681-1263, 682-893, 682-1049, 685-925, 685-957, 685-1055, 686-1051, 701-880, 701-921, 702-944, 703-989, 704-1046, 704-1049, 713-1049, 721-950, 734-1013, 743-1068, 746-1051, 751-935, 753-1003, 753-1048, 753-1051, 753-1053, 753-1068, 754-985, 756-1000, 761-1058, 763-1000, 778-999, 789-1019, 792-1066, 794-1050, 796-1049, 796-1053, 796-1068, 798-1058, 803-1023, 804-1014, 804-1049, 804-1068, 810-1064, 816-1068, 827-1047, 828-1065, 831-1065, 833-1051, 837-1052, 841-1046, 850-1064, 863-1068, 865-1064,
	865-1068, 866-1049, 866-1068, 868-1068, 870-1068, 871-1038, 880-1052, 880-1066, 888-1049, 899-1068, 900-1037, 925-1049, 925-1054, 925-1068, 966-1048
79/ 7504690CB1/ 1854	1-406, 1-1854, 69-620, 69-725, 69-761, 118-364, 118-506, 177-506, 365-506, 501-742, 501-753, 501-767, 517-962, 523-978, 542-729, 544-843, 555-988, 566-1086, 569-819, 574-826, 605-1389, 610-1097, 611-863, 619-1126, 636-897, 647-1063, 656-1122, 656-1388, 657-1350, 664-1389, 665-1444, 669-1086, 694-905, 711-1475, 729-1425, 731-1378, 732-1012, 744-999, 754-1014, 757-1305, 761-1428, 772-1389, 773-1389, 774-1389, 777-1114, 778-1057, 794-1469, 796-1460, 803-1457, 810-1419, 817-1460, 822-1466, 827-1105, 827-1345, 830-1068, 835-1369, 842-1126, 846-1239, 846-1462, 854-1066, 866-1088, 866-1404, 872-1414, 874-1174, 878-1138, 881-1448, 886-1442, 886-1447, 887-1458, 889-1460, 891-1125, 898-1489, 899-1466, 903-1460, 905-1094, 911-1420, 911-1435, 916-1464, 918-1251, 924-1371, 928-1469, 949-1487, 965-1285, 973-1237, 997-1289, 998-1489, 999-1436, 1000-1489, 1009-1466, 1009-1472, 1012-1328, 1014-1148, 1014-1275, 1014-1462, 1016-1325, 1023-1230, 1025-1478, 1030-1261, 1032-1398, 1033-
	1481, 1037-1282, 1037-1448, 1037-1489, 1040-1328, 1042-1256, 1044-1472, 1049-1477, 1053-1226, 1053-1490, 1055-1488, 1056-1475, 1057-1488, 1059-1184, 1063-1234, 1067-1479, 1069-1481, 1085-1479, 1095-1474, 1113-1362, 1142-1408, 1142-1460, 1142-1476, 1142-1489, 1144-1479, 1152-1475, 1154-1277, 1173-1419, 1175-1426, 1187-1477, 1188-1445, 1194-1475, 1196-1476, 1205-1476, 1217-1473, 1218-1479, 1219-1475, 1242-1415, 1247-1469, 1249-1475, 1260-1475, 1270-1479, 1270-1532, 1332-1470, 1341-1477, 1349-1530, 1351-1497, 1376-1486, 1389-1475, 1411-1501, 1635-1840

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
80/ 7504720CB1/ 1332	1-221, 10-71, 10-105, 10-264, 19-1332, 20-98, 30-82, 31-542, 31-593, 32-283, 32-396, 104-348, 104-371, 104-373, 104-452, 104-481, 104-544, 104-545, 104-601, 104-645, 104-646, 104-665, 104-708, 104-719, 104-727, 104-743, 104-787, 105-269, 105-690, 105-700, 107-368, 107-612, 110-386, 112-589, 112-734, 113-743, 114-390, 114-477, 120-218, 120-349, 120-409, 120-665, 122-670, 124-363, 128-754, 133-388, 134-361, 134-826, 135-470, 135-772, 140-385, 143-412, 143-415, 143-800, 144-422, 145-374, 147-385, 151-651, 152-545, 153-681, 154-384, 154-389, 154-695, 154-965, 155-338, 158-425, 160-412, 160-450, 160-642, 161-405, 163-694, 164-456, 165-401, 165-436, 166-400, 167-383, 171-412, 174-468, 174-472, 175-437, 176-646, 183-359, 184-445, 184-456, 185-465, 188-468, 189-501, 189-539, 189-625, 198-476, 206-402, 209-493, 210-479, 210-495, 210-496, 211-470, 211-803, 213-420, 213-464, 213-629, 218-713, 221-483, 221-587, 222-512, 222-707, 224-474, 224-554, 226-694, 231-353, 234-474, 235-949, 236-696,
	237-509, 237-673, 239-496, 239-528, 241-839, 243-506, 243-521, 244-490, 244-506, 245-455, 245-488, 247-584, 250-443, 250-462, 251-765, 252-537, 252-544, 252-714, 259-523, 263-508, 267-565, 268-988, 268-1003, 268-1007, 268-1025, 268-1033, 268-1039, 268-1048, 268-1053, 269-480, 269-542, 269-554, 269-581, 269-721, 270-563, 270-743, 271-486, 271-525, 271-573, 271-762, 272-415, 273-470, 273-949, 276-568, 276-863, 279-622, 281-578, 283-519, 287-533, 288-533, 288-534, 288-592, 288-973, 289-508, 291-525, 291-1047, 292-1004, 298-404, 298-408, 298-409, 298-525, 299-478, 299-546, 300-845, 302-581, 307-567, 310-589, 310-590, 311-546, 313-574, 314-549, 314-862, 316-549, 317-555, 318-600, 318-603, 318-639, 321-596, 322-489, 322-582, 322-635, 322-971, 323-601, 323-611, 324-582, 324-597, 325-584, 326-688, 327-531, 328-584, 328-591, 328-611, 328-920, 330-578, 330-601, 331-581, 331-640, 335-538, 335-841, 337-734, 337-735, 337-834, 337-948, 338-573, 338-579, 338-586, 338-597, 340-560, 340-591,
	343-744, 346-595, 347-775, 355-608, 355-611, 358-626, 358-646, 358-818, 358-869, 361-582, 362-581, 362-876, 363-634, 364-605, 364-606, 364-779, 367-688, 369-623, 369-1284, 370-648, 370-679, 370-1284, 372-642, 372-943, 373-646, 376-592, 378-641, 378-1053, 381-691, 385-623, 385-655, 385-659, 385-683, 388-634, 388-688, 388-907, 393-668, 402-679, 402-754, 403-754, 408-513, 408-594, 408-665, 409-1285, 410-728, 411-682, 411-709, 413-686, 417-1284, 419-678, 419-703, 419-841, 421-682, 422-644, 422-1259, 426-632, 426-638, 426-1284, 427-723, 428-783, 428-1284, 429-1284, 431-698, 432-1284, 433-692, 435-1261, 436-670, 436-698, 436-1054, 437-701, 437-869, 437-992, 437-1070, 441-673, 445-673, 446-1021, 448-660, 448-747, 449-1284, 450-631, 450-704, 450-706, 450-713, 450-715, 450-1001, 450-1094, 450-1109, 454-740, 456-701, 456-716, 456-720, 456-736, 456-747, 457-650, 457-652, 457-759, 457-1020, 460-686, 462-723, 462-1284, 467-1284, 469-748, 470-1244, 471-713, 471-722, 471-728, 472-585, 472-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	695, 472-1060, 477-686, 477-716, 477-758, 477-760, 477-762, 477-782, 478-718, 478-751, 478-1113, 480-716, 480-729, 482-748, 483-661, 483-732, 484-719, 484-732, 484-1033, 485-717, 486-748, 488-764, 489-821, 489-979, 490-667, 490-675, 490-740, 490-750, 491-715, 491-772, 491-783, 491-972, 492-804, 493-1078, 496-786, 499-979, 503-679, 505-754, 505-810, 507-681, 512-786, 513-809, 513-1302, 514-618, 514-786, 514-792, 514-832, 515-686, 515-1149, 518-767, 518-1109, 518-1118, 520-784, 520-786, 520-1123, 522-1250, 523-1013, 524-652, 524-794, 525-690, 525-785, 525-786, 525-794, 526-753, 526-814, 526-859, 526-984, 526-1060, 527-698, 527-877, 528-729, 528-773, 528-784, 529-770, 529-771, 529-930, 531-839, 533-745, 533-800, 533-830, 533-1330, 534-753, 534-755, 534-764, 534-777, 534-797, 534-812, 534-813, 534-814, 534-973, 535-1267, 536-1256, 538-784, 538-797, 539-885, 540-1233, 541-824, 543-1328, 544-805, 548-659, 548-807, 548-850, 548-851, 548-1113, 548-1321, 549-1252, 549-1255, 551-804,
	551-1263, 555-833, 555-1007, 556-819, 557-1265, 558-792, 561-1019, 563-846, 564-1321, 565-895, 566-825, 566-1171, 568-830, 568-832, 568-852, 569-916, 570-720, 570-742, 570-822, 570-828, 570-834, 570-836, 570-877, 574-841, 574-842, 576-1269, 577-857, 578-817, 579-1085, 580-872, 581-1047, 582-815, 582-1279, 583-869, 585-778, 585-807, 585-848, 585-871, 586-837, 586-851, 587-902, 588-823, 588-825, 588-862, 588-869, 588-1262, 589-825, 592-840, 592-1190, 593-764, 593-790, 593-800, 593-1002, 593-1320, 594-757, 594-775, 594-786, 594-810, 594-825, 594-870, 594-887, 594-895, 594-1069, 594-1317, 595-755, 595-819, 595-820, 595-868, 596-859, 596-1321, 597-1033, 600-1321, 602-1248, 604-803, 604-963, 604-996, 604-1240, 605-869, 605-897, 605-898, 605-1132, 605-1279, 609-1261, 609-1266, 609-1310, 611-1191, 614-1276, 615-905, 616-1286, 616-1328, 619-842, 620-914, 620-1262, 620-1321, 621-889, 622-775, 622-825, 622-910, 625-892, 626-896, 626-898, 626-976, 626-978, 626-1238, 628-1075,
	629-852, 629-875, 629-910, 629-920, 630-1052, 631-1321, 634-874, 634-912, 634-966, 635-882, 635-898, 637-1187, 639-1315, 640-1321, 641-794, 641-905, 643-881, 643-882, 645-1315, 645-1328, 646-1263, 646-1265, 648-882, 650-1155, 651-861, 651-942, 651-947, 651-949, 651-960, 651-1207, 653-872, 654-1282, 654-1331, 654-1332, 655-820, 655-903, 655-1264, 655-1313, 655-1315, 655-1321, 656-1308, 663-946, 663-1268, 664-909, 664-915, 664-916, 664-917, 665-867, 665-916, 665-942, 666-878, 666-1121, 667-1328, 668-913, 668-926, 668-1052, 668-1090, 669-781, 669-938, 669-953, 670-806, 670-855, 670-861, 670-875, 670-895, 670-898, 670-1033, 670-1209, 670-1212, 671-908, 671-945, 672-919, 673-940, 674-1155, 676-982, 676-1010, 678-915, 686-935, 686-962, 686-984, 686-1000, 687-1120, 688-926, 689-971, 690-942, 693-919, 693-920, 693-1321, 694-945, 694-960, 694-969, 695-901, 695-1147, 695-1269, 698-951, 699-1063, 702-952, 703-1031, 703-1253, 704-916, 705-882, 705-905, 705-1303, 705-1310, 705-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1313, 710-1268, 711-867, 711-996, 712-743, 712-942, 712-984, 712-987, 712-1003, 712-1127, 714-1026, 714-1230, 718-963, 718-998, 718-1009, 719-1141, 720-968, 720-984, 720-1019, 721-934, 721-1100, 725-1085, 725-1205, 726-1307, 727-1261, 730-994, 730-1285, 732-1296, 734-1286, 735-972, 736-1290, 738-1224, 738-1332, 739-860, 739-946, 739-958, 739-1332, 740-996, 743-1268, 745-1332, 746-1313, 746-1321, 750-989, 750-990, 750-1011, 750-1041, 750-1325, 750-1332, 751-905, 755-961, 755-1006, 755-1332, 756-1315, 758-1311, 761-988, 761-991, 761-1030, 761-1040, 761-1269, 761-1332, 762-1332, 763-1315, 764-1332, 765-977, 766-1332, 767-1311, 768-1224, 768-1321, 771-953, 771-1015, 771-1020, 771-1033, 771-1042, 771-1048, 771-1056, 772-897, 772-989, 772-991, 772-1049, 772-1331, 772-1332, 773-1315, 774-1049, 774-1060, 774-1316, 776-927, 776-1010, 776-1318, 781-1024, 781-1035, 781-1262, 781-1332, 782-1000, 782-1039, 783-1307, 783-1332, 784-965, 784-1037, 784-1052, 784-1057, 784-1157, 786-
	1044, 786-1085, 786-1286, 787-1332, 791-1332, 793-1102, 794-1053, 795-1046, 795-1059, 795-1072, 795-1088, 795-1332, 796-1020, 796-1028, 797-972, 799-1079, 799-1285, 800-1281, 801-1015, 801-1035, 801-1267, 801-1332, 802-941, 802-1018, 802-1031, 802-1198, 802-1238, 803-1332, 804-1049, 807-1320, 808-1332, 810-1327, 812-1069, 812-1076, 813-1244, 813-1332, 815-1323, 815-1332, 816-1055, 816-1274, 816-1332, 817-1294, 819-1058, 820-1209, 821-977, 821-1069, 821-1085, 821-1131, 821-1240, 822-1032, 822-1037, 822-1054, 822-1088, 822-1253, 822-1332, 823-969, 823-1073, 824-1239, 825-1317, 826-1020, 826-1025, 826-1058, 826-1079, 826-1087, 826-1332, 827-1062, 827-1083, 827-1332, 830-1088, 830-1238, 830-1332, 834-1115, 836-1332, 837-1083, 838-1332, 839-1098, 839-1116, 840-984, 840-1089, 840-1176, 840-1329, 841-1102, 841-1112, 841-1115, 841-1125, 841-1142, 842-1122, 843-1061, 843-1062, 843-1132, 843-1199, 843-1323, 844-1032, 844-1098, 844-1118, 845-1138, 846-1094, 846-1109, 846-1235, 846-1326,
	847-1123, 847-1133, 848-1124, 849-1323, 849-1332, 851-1093, 851-1109, 851-1118, 853-1032, 853-1075, 853-1109, 853-1196, 853-1216, 855-1039, 855-1134, 855-1202, 856-1109, 856-1311, 856-1322, 857-1090, 857-1118, 858-1332, 860-1315, 860-1316, 860-1319, 860-1328, 860-1332, 861-1097, 862-1327, 862-1332, 863-1167, 863-1185, 863-1327, 864-1316, 864-1323, 864-1332, 865-1321, 865-1323, 866-1066, 866-1325, 867-1323, 869-1321, 869-1323, 870-1324, 870-1328, 871-986, 871-1087, 872-1112, 872-1167, 872-1324, 873-995, 873-1322, 873-1332, 874-1324, 880-1122, 880-1323, 881-1281, 881-1320, 881-1332, 882-990, 882-1320, 882-1325, 882-1332, 883-1252, 884-1324, 885-1076, 885-1328, 885-1332, 886-1095, 889-1332, 890-1331, 891-1156, 891-1332, 892-1086, 892-1323, 893-1123, 893-1142, 893-1207, 895-1151, 895-1315, 895-1322, 896-1187, 896-1314, 896-1323, 898-1325, 899-1127, 899-1320, 900-1332, 901-1321, 901-1325, 902-1230, 902-1323, 904-1224, 904-1323, 904-1328, 905-1320, 906-1323, 908-1323, 908-1324, 908-1327,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	909-1117, 909-1320, 909-1323, 909-1332, 910-1324, 910-1332, 911-1188, 911-1311, 912-1320, 912-1325, 913-1158, 913-1326, 913-1332, 914-1158, 914-1325, 916-1154, 916-1169, 916-1315, 916-1323, 916-1324, 917-1324, 917-1325, 918-1330, 919-1131, 919-1147, 919-1167, 919-1178, 919-1184, 919-1324, 919-1325, 920-1323, 921-1163, 921-1175, 921-1182, 921-1323, 922-1323, 923-1324, 923-1332, 924-1324, 927-1324, 930-1313, 930-1324, 931-1323, 931-1332, 932-1332, 933-1160, 934-1151, 934-1205, 936-1332, 938-1309, 938-1316, 941-1332, 943-1332, 944-1281, 945-1319, 949-1310, 949-1322, 949-1325, 950-1040, 950-1208, 950-1319, 950-1323, 951-1321, 951-1323, 952-1212, 952-1323, 953-1277, 953-1324, 954-1321, 954-1323, 954-1324, 955-1323, 956-1323, 957-1327, 958-1326, 959-1218, 961-1210, 961-1252, 961-1322, 962-1285, 962-1324, 964-1259, 964-1331, 966-1229, 968-1225, 968-1320, 968-1323, 969-1324, 970-1324, 970-1327, 971-1226, 971-1257, 971-1266, 971-1324, 972-1320, 973-1129, 973-1323, 974-1323, 975-1213, 976-1259,
	976-1263, 976-1272, 976-1323, 977-1318, 978-1233, 979-1324, 980-1323, 981-1208, 981-1224, 981-1323, 983-1275, 985-1320, 988-1323, 989-1231, 989-1323, 991-1199, 992-1323, 992-1328, 993-1332, 995-1294, 996-1331, 996-1332, 997-1332, 998-1317, 999-1324, 1002-1257, 1003-1256, 1003-1326, 1006-1242, 1006-1257, 1006-1263, 1006-1269, 1006-1320, 1006-1321, 1006-1323, 1006-1324, 1007-1270, 1007-1323, 1008-1323, 1008-1326, 1009-1235, 1009-1287, 1009-1317, 1010-1318, 1011-1322, 1011-1324, 1011-1328, 1012-1323, 1013-1324, 1017-1269, 1018-1332, 1021-1293, 1022-1322, 1022-1323, 1023-1326, 1025-1323, 1027-1324, 1028-1326, 1029-1332, 1030-1330, 1033-1323, 1035-1329, 1036-1313, 1037-1323, 1038-1323, 1038-1326, 1039-1218, 1039-1313, 1039-1325, 1041-1317, 1042-1315, 1043-1323, 1045-1323, 1045-1324, 1046-1303, 1046-1323, 1050-1272, 1053-1252, 1054-1228, 1056-1317, 1056-1320, 1056-1332, 1063-1283, 1066-1312, 1066-1327, 1075-1320, 1077-1332, 1078-1323, 1079-
	1296, 1083-1268, 1083-1294, 1083-1315, 1083-1332, 1086-1332, 1094-1332, 1099-1250, 1099-1332, 1102-1324, 1115-1320, 1115-1323, 1115-1324, 1116-1332, 1117-1317, 1118-1323, 1118-1329, 1123-1326, 1125-1323, 1128-1325, 1129-1332, 1130-1323, 1131-1322, 1132-1332, 1134-1332, 1143-1329, 1146-1322, 1172-1320, 1174-1323, 1177-1323, 1186-1332, 1196-1294, 1200-1332, 1202-1325, 1202-1332, 1215-1332, 1216-1332, 1237-1323, 1237-1332, 1239-1332, 1256-1332, 1259-1327, 1259-1332, 1261-1332
81/ 7504722CB1/ 1435	1-276, 2-1435, 12-210, 12-241, 12-247, 12-297, 12-303, 12-332, 14-259, 14-264, 14-295, 19-135, 19-245, 19-259, 23-288, 23-291, 23-308, 24-295, 26-295, 30-281, 30-290, 30-295, 31-244, 31-295, 32-276, 33-165, 33-261, 33-282, 33-362, 34-302, 37-301, 40-292, 47-263, 48-326, 54-682, 56-263, 56-288, 56-295, 56-297, 56-305, 56-308, 56-318, 57-314, 57-397, 57-670, 58-321, 60-179, 60-343, 60-353, 62-210, 62-295, 62-312, 63-295, 69-193, 69-208, 69-374, 71-390, 71-617, 72-295, 72-318, 72-319, 72-323, 72-331, 73-272, 73-286, 73-346, 73-363, 74-302, 76-239, 76-325, 76-349, 77-196, 77-211, 77-265, 77-287, 77-293, 77-295, 77-310, 77-353, 77-355, 77-358, 78-268, 78-295, 78-302, 78-325, 78-342, 78-357, 78-358, 78-363, 78-373, 78-382, 78-387, 78-397, 79-341, 79-397, 80-236, 80-323, 80-356, 80-361, 81-295, 81-355, 81-362, 83-295, 83-342, 83-370, 86-329, 87-332, 87-355, 88-352, 90-177, 90-381, 91-357, 92-268, 93-334, 94-330, 94-375, 94-397, 95-371, 95-373, 96-355, 96-720, 99-336, 103-329, 105-367, 105-387,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	107-397, 108-342, 108-383, 109-381, 115-364, 122-384, 123-396, 124-289, 124-371, 126-369, 132-624, 145-363, 178-374, 190-397, 209-397, 221-468, 236-339, 239-397, 292-397, 295-397, 393-562, 397-481, 402-641, 508-618, 508-721, 520-1041, 527-1058, 677-1265, 677-1383, 697-995, 718-975, 773-1317, 798-984, 850-1126, 888-1134, 1082-1332, 1098-1313, 1100-1386, 1109-1393, 1111-1375, 1117-1389, 1131-1364, 1139-1385, 1144-1375, 1156-1246, 1173-1317, 1200-1395
82/ 7504733CB1/ 554	1-543, 153-553, 157-407, 157-422, 157-442, 157-459, 157-526, 157-534, 157-542, 157-551, 157-554, 159-513, 160-554, 161-392, 202-554, 280-553, 378-547
83/ 7507100CB1/ 2365	1-2365, 165-521, 165-535, 171-535, 174-325, 174-393, 178-541, 178-706, 178-830, 178-991, 178-1126, 296-380, 854-979, 876-1402, 1117-1354, 1184-1510, 1211-1505, 1236-1760, 1271-1579, 1276-1884, 1300-1591, 1368-1506, 1368-1619, 1368-1760, 1368-1804, 1368-1805, 1368-1823, 1368-1887, 1368-2027, 1372-1643, 1388-1757, 1420-1982, 1488-2130, 1507-1830, 1507-2128, 1567-1834, 1590-2153, 1590-2164, 1596-2165, 1612-1932, 1614-2119, 1614-2128, 1624-1903, 1682-2145, 1701-1885, 1709-1938, 1709-1956, 1742-2096, 1776-2089, 1802-2282, 1822-2086, 1826-2131, 1834-2033, 1864-2189, 1865-2361, 1869-2104, 1877-2241, 1881-2173, 1891-2222, 1935-2162, 1937-2116, 1946-2134, 1978-2177, 1979-2176, 2036-2180, 2082-2361, 2201-2365
84/ 7503330CB1/ 703	1-115, 13-140, 13-684, 26-99, 72-140, 132-609, 132-623, 132-648, 132-658, 132-659, 132-683, 132-684, 134-539, 142-519, 147-297, 148-683, 158-534, 164-687, 168-669, 179-389, 181-509, 195-683, 219-473, 223-703, 239-653, 275-701, 277-551, 381-596, 464-700, 464-703, 607-696
85/ 7504519CB1/ 3871	1-3171, 39-3187, 51-612, 51-812, 51-852, 52-807, 326-1049, 481-1332, 527-1150, 528-1322, 560-1122, 616-1508, 631-1304, 632-1480, 651-1036, 651-1138, 651-1284, 651-1296, 651-1298, 651-1321, 651-1344, 651-1353, 651-1368, 651-1377, 651-1391, 651-1397, 651-1410, 651-1437, 651-1440, 651-1447, 651-1449, 651-1453, 651-1468, 651-1481, 651-1488, 651-1494, 651-1495, 651-1498, 651-1499, 651-1511, 651-1514, 651-1515, 651-1518, 651-1522, 651-1523, 651-1526, 651-1528, 651-1533, 651-1537, 651-1539, 651-1543, 651-1547, 651-1548, 651-1552, 651-1553, 651-1562, 651-1564, 651-1565, 651-1567, 651-1579, 651-1586, 651-1608, 651-1656, 651-1658, 653-1517, 653-1577, 663-738, 716-1052, 720-1481, 747-1462, 750-1472, 758-1394, 783-1493, 810-1408, 832-1608, 835-1536, 840-1341, 901-1145, 905-1063, 933-1700, 944-1099, 985-1530, 1052-1732, 1060-1693, 1106-1339, 1106-1353, 1106-1572, 1106-1631, 1304-2180, 1389-2180, 1452-2180, 1466-2180, 1535-1735, 1543-2180, 1652-2308, 1705-2291, 1720-2333, 1727-2347, 1728-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	2256, 1746-2396, 1754-2215, 1775-2423, 1776-2262, 1785-2559, 1787-2302, 1789-2382, 1799-2383, 1806-2360, 1809-2347, 1834-2450, 1838-2395, 1841-2349, 1846-2103, 1848-2470, 1856-2418, 1856-2502, 1857-2389, 1858-2565, 1860-2413, 1861-2363, 1861-2389, 1861-2659, 1868-2444, 1877-2502, 1882-2415, 1885-2497, 1885-2520, 1886-2464, 1891-2470, 1891-2506, 1899-2536, 1900-2683, 1906-2441, 1909-2437, 1909-2461, 1911-2403, 1911-2539, 1914-2526, 1915-2453, 1916-2392, 1916-2445, 1917-2566, 1922-2240, 1924-2520, 1929-2491, 1933-2553, 1943-2180, 1945-2312, 1945-2502, 1947-2470, 1948-2285, 1948-2596, 1955-2497, 1962-2511, 1963-2490, 1969-2517, 1972-2515, 1980-2614, 1987-2090, 1992-2708, 1995-2679, 1997-2629, 1998-2596, 2001-2595, 2002-2572, 2003-2610, 2007-2224, 2013-2638, 2020-2541, 2024-2246, 2027-2622, 2042-2622, 2056-2598, 2058-2386, 2058-2591, 2066-2608, 2074-2616, 2086-2685, 2089-2699, 2104-2698, 2127-2248, 2127-2694, 2132-2512, 2133-
	2681, 2133-2695, 2142-2763, 2149-2362, 2154-2446, 2154-2605, 2154-2660, 2159-2290, 2160-2685, 2161-2660, 2162-2681, 2164-2687, 2165-2660, 2166-2761, 2169-2723, 2171-2796, 2177-2580, 2181-2392, 2181-2637, 2181-2639, 2182-2729, 2184-2697, 2187-2673, 2188-2763, 2194-2690, 2196-2437, 2204-2779, 2207-2775, 2215-2593, 2216-2595, 2237-2783, 2239-2700, 2247-2523, 2248-2380, 2258-2819, 2269-2555, 2269-2836, 2271-2706, 2278-2522, 2279-2872, 2282-2365, 2282-2572, 2282-2760, 2282-2790, 2282-2869, 2282-2871, 2282-2876, 2282-2891, 2282-2920, 2282-3099, 2303-2646, 2322-2837, 2333-2557, 2336-2717, 2345-2557, 2414-2704, 2425-3085, 2457-2709, 2457-2757, 2457-2992, 2480-2728, 2481-2581, 2521-2624, 2534-3167, 2546-2692, 2576-2662, 2581-3147, 2587-2876, 2594-3160, 2603-2872, 2621-2928, 2690-2947, 2700-2966, 2730-3162, 2731-3228, 2738-3022, 2744-3225, 2745-2947, 2750-3177, 2751-3202, 2751-3227, 2756-3871, 2781-2947, 2786-2944, 2787-3228, 2789-3156, 2804-2938, 2828-2966, 2831-3000, 2837-3208, 2886-2947, 2886-3203
86/ 7504705CB1/ 695	1-250, 1-634, 103-244, 106-353, 112-230, 116-366, 117-348, 117-649, 118-251, 121-444, 122-239, 123-251, 124-416, 125-215, 125-229, 251-487, 253-695, 254-466, 255-484, 255-515, 258-456, 258-503, 268-478, 268-481, 278-482, 282-481, 371-645, 379-648, 383-641, 407-634, 408-635, 432-633, 432-657, 433-560, 433-656, 436-679, 438-678, 438-679, 440-645, 452-635, 467-646, 472-652, 489-634, 567-695
87/ 7504738CB1/ 536	1-132, 1-142, 1-241, 1-246, 1-536, 2-135, 5-110, 5-115, 5-127, 5-142, 5-273, 5-304, 15-142, 49-142, 78-531, 97-341, 125-364, 125-468, 125-531, 142-371, 143-292, 143-344, 143-351, 143-353, 143-393, 143-420, 143-518, 143-536, 144-410, 145-398, 147-386, 148-301, 148-410, 150-524, 153-384, 163-402, 163-410, 164-317, 164-411, 164-485, 170-524, 174-437, 179-534, 184-524, 185-518, 186-524, 188-417, 188-526, 200-402, 200-466, 200-482, 200-515, 216-451, 220-472, 269-531, 275-524, 278-528, 279-534, 296-536, 312-536, 325-515, 329-536, 330-524, 333-536, 335-532, 337-528, 354-528, 367-477, 367-532, 367-536, 368-532, 368-536, 372-536, 377-532, 384-532, 395-528, 410-524, 411-528, 412-532, 420-528, 423-536, 434-536, 439-528, 443-535, 444-528, 448-528, 457-528
88/ 7510280CB1/ 549	1-268, 1-549, 10-399, 12-530, 16-399, 23-344, 28-150, 28-271, 211-499

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
89/ 7503700CB1/ 1888	1-1868, 404-692, 416-828, 422-698, 429-676, 530-747, 530-765, 530-1056, 586-858, 624-835, 744-1052, 761-1021, 780-1070, 897-1056, 1060-1291, 1090-1296, 1090-1809, 1137-1431, 1137-1806, 1195-1425, 1209-1434, 1260-1846, 1265-1868, 1268-1857, 1342-1634, 1343-1568, 1371-1868, 1397-1843, 1400-1868, 1403-1888, 1405-1862, 1409-1864, 1439-1862, 1441-1866, 1442-1818, 1450-1867, 1452-1860, 1457-1865, 1474-1865, 1490-1868, 1496-1860, 1498-1868, 1500-1860, 1501-1860, 1503-1865, 1505-1860, 1506-1865, 1511-1860, 1512-1860, 1524-1871, 1536-1872, 1539-1862, 1541-1859, 1545-1856, 1545-1858, 1545-1864, 1545-1888, 1548-1860, 1549-1860, 1549-1864, 1554-1864, 1556-1838, 1556-1858, 1556-1864, 1557-1838, 1557-1856, 1557-1858, 1557-1864, 1575-1860, 1585-1865, 1606-1868, 1682-1845, 1696-1888, 1773-1860, 1805-1868
90/ 7504685CB1/ 1050	1-260, 8-265, 8-283, 11-231, 15-212, 15-488, 21-490, 22-255, 25-926, 33-275, 36-207, 36-296, 48-310, 50-345, 51-329, 60-340, 61-484, 69-326, 84-481, 84-496, 108-234, 115-487, 160-488, 164-530, 282-395, 301-876, 318-556, 328-422, 337-484, 370-984, 389-649, 389-967, 395-488, 395-989, 406-513, 413-679, 444-511, 447-727, 452-927, 508-752, 508-754, 515-764, 522-764, 566-859, 582-920, 589-831, 589-846, 597-1050, 601-848, 622-865, 666-926, 679-893, 717-1012
91/ 7506844CB1/ 899	1-137, 1-191, 1-201, 1-878, 1-894, 46-195, 64-755, 64-791, 197-795, 197-891, 200-672, 214-743, 227-776, 231-803, 239-762, 240-795, 245-532, 252-531, 270-819, 289-865, 299-873, 317-788, 318-870, 319-573, 335-858, 335-893, 342-840, 358-899, 360-853, 360-884, 366-652, 371-640, 381-643, 382-534, 386-520, 392-705, 393-604, 405-648, 405-662, 410-881, 414-788, 414-899, 416-837, 416-899, 420-881, 423-674, 423-842, 423-881, 424-898, 426-853, 430-881, 430-883, 436-539, 438-673, 438-881, 439-888, 444-715, 447-881, 452-689, 452-714, 452-796, 454-886, 456-661, 457-881, 458-656, 458-741, 462-882, 464-881, 468-881, 475-881, 479-880, 480-677, 480-881, 481-733, 481-881, 481-883, 487-899, 492-882, 494-660, 496-754, 504-779, 507-899, 508-880, 515-762, 517-881, 517-882, 518-819, 521-881, 521-884, 522-883, 524-881, 540-883, 540-899, 541-800, 545-881, 546-888, 552-899, 565-714, 579-899, 580-899, 584-899, 587-881, 600-899, 618-896, 622-871, 622-879, 625-757, 625-882, 627-868, 632-821, 633-803, 641-896, 708-839, 712-881, 727-898, 751-881, 770-879
92/ 7510259CB1/ 2529	1-147, 1-173, 1-573, 1-2514, 3-573, 75-573, 182-573, 1189-1526, 1189-1540, 1189-1616, 1189-1649, 1189-1653, 1189-1704, 1189-1735, 1189-1793, 1189-1916, 1189-1958, 1189-1979, 1194-1953, 1195-1828, 1353-2121, 1449-2226, 1637-2393, 1650-2511, 1683-2393, 1832-2511, 1918-2529, 1955-2522, 2004-2517, 2117-2524

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
93/ 7510444CB1/ 2268	1-162, 1-2268, 22-806, 22-815, 160-557, 328-1237, 465-1045, 472-1229, 492-977, 533-806, 535-1120, 662-1249, 684-1144, 734-1144, 743-1145, 746-1309, 749-1145, 777-1108, 817-1145, 861-1337, 889-1145, 890-1145, 932-1459, 956-1251, 957-1583, 964-1323, 1021-1585, 1043-1185, 1052-1898, 1060-1743, 1125-1792, 1143-1559, 1143-1729, 1143-1750, 1143-1758, 1146-1537, 1170-1833, 1212-1874, 1227-1792, 1240-1770, 1260-1723, 1268-1924, 1270-2000, 1292-1745, 1297-2148, 1334-2006, 1338-1947, 1348-2148, 1357-2138, 1375-2148, 1381-2148, 1393-2035, 1402-2148, 1406-2003, 1428-1749, 1433-1985, 1446-2003, 1446-2148, 1471-2191, 1486-1745, 1499-2232, 1505-2148, 1528-1853, 1531-2245, 1571-2261, 1576-2144, 1578-2191, 1584-2234, 1632-2191, 1643-2268, 1679-2184, 1695-2266, 1727-2268, 1783-2172, 1789-2064, 1800-2039, 1853-2238
94/ 7510494CB1/ 1819	1-670, 1-688, 1-1819, 170-748, 183-765, 213-479, 218-789, 249-772, 266-809, 268-379, 306-647, 361-650, 364-1019, 376-923, 377-1003, 384-594, 385-664, 386-902, 410-795, 419-647, 421-630, 421-657, 421-659, 421-664, 421-672, 423-1310, 427-661, 442-1001, 449-918, 449-948, 449-989, 451-709, 458-994, 461-1086, 462-1181, 476-1019, 482-770, 491-633, 493-1019, 505-774, 505-776, 507-1310, 510-1315, 512-776, 515-761, 515-768, 518-1230, 523-759, 524-1101, 533-861, 535-808, 542-759, 542-871, 542-1264, 543-805, 543-819, 544-1517, 550-1310, 551-1310, 554-760, 560-1310, 562-1338, 563-1512, 564-1517, 569-1518, 574-855, 574-861, 576-875, 582-1097, 589-1310, 592-1168, 597-1517, 597-1518, 601-1310, 603-955, 610-1517, 611-988, 616-1138, 620-927, 620-1310, 627-1089, 629-893, 635-828, 635-1118, 638-1205, 642-1165, 642-1512, 648-892, 648-895, 651-899, 653-858, 656-1513, 668-890, 668-1160, 670-1186, 671-1073, 672-1031, 682-1310, 686-938, 689-1033, 692-1067, 694-967, 696-1204, 697-
	967, 699-943, 699-1295, 703-1132, 712-1222, 714-1513, 714-1518, 715-967, 716-1518, 717-1141, 723-1364, 726-1032, 726-1035, 729-1305, 730-978, 731-1029, 731-1058, 736-963, 739-1255, 740-1058, 742-1296, 746-1021, 746-1077, 756-965, 764-1306, 774-1049, 783-1050, 792-1111, 793-1447, 793-1517, 797-1518, 798-1471, 799-1437, 800-1024, 800-1026, 818-1080, 819-1172, 821-1081, 824-1367, 824-1512, 825-1122, 828-1518, 834-1078, 838-1350, 839-1446, 852-1433, 857-1107, 862-1104, 880-1087, 882-1161, 884-1118, 886-1157, 889-1163, 898-1104, 904-1446, 905-1117, 905-1124, 916-1126, 916-1212, 916-1465, 920-1177, 921-1442, 922-1219, 922-1441, 922-1510, 926-1170, 926-1180, 928-1221, 944-1230, 947-1640, 948-1626, 949-1227, 952-1158, 952-1195, 952-1516, 956-1210, 956-1521, 963-1202, 968-1243, 968-1258, 968-1266, 988-1330, 989-1228, 995-1302, 998-1206, 998-1313, 1004-1197, 1006-1423, 1009-1758, 1010-1226, 1010-1552, 1012-1226, 1015-1511, 1021-1276, 1025-1360, 1027-1537, 1032-1277, 1035-1345,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1036-1321, 1037-1335, 1039-1299, 1043-1281, 1045-1521, 1048-1192, 1059-1319, 1061-1317, 1062-1346, 1062-1510, 1065-1223, 1065-1551, 1068-1361, 1071-1716, 1075-1589, 1077-1338, 1084-1337, 1085-1383, 1098-1337, 1101-1382, 1104-1464, 1108-1386, 1114-1367, 1118-1610, 1119-1392, 1125-1324, 1141-1413, 1141-1419, 1141-1446, 1142-1430, 1148-1410, 1153-1406, 1155-1402, 1160-1396, 1160-1412, 1160-1421, 1162-1561, 1164-1424, 1166-1424, 1168-1436, 1169-1434, 1184-1436, 1201-1475, 1203-1454, 1204-1435, 1220-1508, 1227-1508, 1228-1433, 1232-1489, 1233-1517, 1283-1522, 1295-1543, 1513-1758, 1526-1812, 1538-1802, 1557-1798, 1560-1790, 1565-1793, 1577-1816, 1647-1818
95/ 6486485CB1/ 1711	1-1001, 915-1518, 915-1544, 927-1095, 935-1180, 946-1168, 955-1324, 1287-1453, 1345-1410, 1424-1711, 1451-1610, 1467-1711, 1469-1711, 1504-1711, 1651-1711
	1036-1321, 1037-1335, 1039-1299, 1043-1281, 1045-1521, 1048-1192, 1059-1319, 1061-1317, 1062-1346, 1062-1510, 1065-1223, 1065-1551, 1068-1361, 1071-1716, 1075-1589, 1077-1338, 1084-1337, 1085-1383, 1098-1337, 1101-1382, 1104-1464, 1108-1386, 1114-1367, 1118-1610, 1119-1392, 1125-1324, 1141-1413, 1141-1419, 1141-1446, 1142-1430, 1148-1410, 1153-1406, 1155-1402, 1160-1396, 1160-1412, 1160-1421, 1162-1561, 1164-1424, 1166-1424, 1168-1436, 1169-1434, 1184-1436, 1201-1475, 1203-1454, 1204-1435, 1220-1508, 1227-1508, 1228-1433, 1232-1489, 1233-1517, 1283-1522, 1295-1543, 1513-1758, 1526-1812, 1538-1802, 1557-1798, 1560-1790, 1565-1793, 1577-1816, 1647-1818
96/ 7503772CB1/ 5096	1-640, 1-5096, 2-5096, 449-1261, 452-1261, 480-1261, 482-1261, 490-1261, 547-1261, 606-1260, 620-1261, 669-1261, 670-1261, 850-1261, 863-1588, 956-1261, 970-1261, 1185-1799, 1199-1640, 1294-1695, 1337-1997, 1355-2041, 1743-2441, 2074-2478, 2083-2216, 2083-2351, 2083-2591, 2083-2602, 2083-2652, 2083-2763, 2103-2477, 2103-2478, 2129-2478, 2147-2756, 2206-2876, 2235-3032, 2239-2488, 2271-2490, 2368-2611, 2434-3066, 2528-2732, 2761-2884, 2893-3010, 3014-3702, 3178-3271, 3324-3948, 3341-3988, 3341-4066, 3341-4159, 3345-3951, 3345-4048, 3349-4159, 3389-3546, 3397-4031, 3402-4001, 3403-3582, 3403-3627, 3403-3696, 3403-3779, 3403-3875, 3403-3947, 3403-4011, 3403-4128, 3403-4135, 3403-4137, 3403-4159, 3403-4171, 3403-4179, 3403-4186, 3404-4159, 3414-4375, 3420-3988, 3430-4334, 3438-4116, 3441-4091, 3452-4064, 3458-4097, 3465-4036, 3497-4154, 3507-4094, 3509-4260, 3519-4095, 3546-4392, 3549-4419, 3551-3796, 3574-4147, 3605-4145, 3611-4459,
	3619-3728, 3637-4306, 3642-4188, 3644-4314, 3656-4237, 3691-3931, 3704-3884, 3713-4334, 3737-4420, 3765-4081, 3790-4358, 3814-4342, 3827-4072, 4076-4714, 4091-4763, 4124-4418, 4330-4813, 4357-4813, 4461-4737, 4476-4743, 4476-4759, 4483-4813, 4503-4759, 4601-4813, 4608-4813, 4683-4813

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
97/ 7503773CB1/ 5049	1-640, 1-5049, 2-5049, 449-1261, 452-1261, 480-1261, 482-1261, 485-1261, 547-1261, 606-1260, 620-1261, 662-1261, 669-1261, 850-1261, 863-1588, 956-1261, 970-1261, 1185-1799, 1199-1640, 1294-1695, 1337-1997, 1355-2041, 1743-2441, 2074-2478, 2083-2216, 2083-2351, 2083-2591, 2083-2602, 2083-2652, 2083-2763, 2103-2477, 2103-2478, 2129-2478, 2147-2756, 2206-2876, 2235-3032, 2239-2488, 2271-2490, 2368-2611, 2434-3066, 2528-2732, 2700-3225, 2704-3328, 2715-3319, 2761-2884, 2871-3508, 2895-3594, 2920-3625, 2939-3339, 2957-3490, 2964-3574, 3022-3835, 3028-3712, 3033-3481, 3033-3590, 3034-3533, 3040-3581, 3045-3449, 3053-3792, 3054-3704, 3058-3721, 3079-3571, 3100-3626, 3105-3712, 3122-3710, 3154-3850, 3165-3776, 3169-3732, 3182-3927, 3186-3813, 3187-3737, 3194-3928, 3198-3865, 3213-3758, 3225-3915, 3242-3932, 3245-3682, 3253-3814, 3257-3527, 3263-4030, 3267-3871, 3303-4048, 3325-3980, 3603-4456, 3650-4308, 3662-4393, 3705-4327, 3738-4376,
	3771-4237, 3773-4013, 3786-3966, 3815-4460, 3820-4199, 3853-4505, 3876-4477, 3934-4357, 3939-4700, 4048-4549, 4048-4664, 4048-4714, 4048-4763, 4048-4766, 4048-4767, 4049-4763, 4053-4313, 4055-4694, 4055-4766, 4059-4766, 4068-4766, 4069-4766, 4075-4766, 4077-4676, 4099-4766, 4100-4766, 4106-4766, 4109-4766, 4115-4766, 4117-4766, 4124-4469, 4132-4372, 4132-4766, 4134-4766, 4135-4766, 4138-4766, 4139-4766, 4139-4767, 4141-4713, 4142-4756, 4145-4364, 4149-4761, 4151-4384, 4157-4763, 4161-4699, 4167-4328, 4169-4763, 4171-4761, 4173-4766, 4179-4766, 4180-4632, 4190-4766, 4195-4766, 4198-4756, 4198-4766, 4209-4766, 4210-4766, 4212-4757, 4214-4766, 4219-4766, 4237-4766, 4247-4565, 4248-4766, 4259-4807, 4267-4766, 4268-4728, 4269-4756, 4276-4766, 4278-4521, 4278-4766, 4287-4761, 4289-4766, 4303-4766, 4306-4767, 4310-4766, 4311-4766, 4320-4766, 4324-4604, 4325-4766, 4331-4598, 4336-4766, 4338-4766, 4352-4766, 4353-4766, 4362-4766,
	4379-4766, 4385-4766, 4390-4766, 4399-4766, 4405-4766, 4414-4690, 4418-4766, 4419-4677, 4426-4766, 4430-4766, 4436-4766, 4440-4766, 4447-4766, 4464-4766, 4470-4703, 4475-4766, 4480-4683, 4480-4766, 4489-4790, 4505-4791, 4511-4766, 4522-4766, 4526-4761, 4528-4766, 4532-4766, 4550-4779, 4554-4766, 4561-4766, 4561-4819, 4561-5049, 4627-4749, 4636-4766, 4651-4745, 4710-4939, 4733-5020, 4752-4914, 4855-5049
98/ 7504698CB1/ 1086	1-133, 1-239, 1-252, 1-1042, 1-1046, 5-216, 11-286, 11-303, 18-288, 18-295, 20-269, 20-301, 27-228, 27-429, 28-159, 28-221, 28-290, 28-292, 28-294, 28-295, 28-305, 28-313, 28-334, 31-273, 31-277, 31-295, 31-303, 32-310, 32-317, 33-188, 33-278, 33-281, 33-319, 34-176, 34-262, 34-284, 34-307, 34-310, 34-318, 34-342, 34-357, 35-262, 35-295, 35-297, 35-381, 36-276, 37-244, 37-262, 37-293, 37-298, 37-300, 37-303, 37-307, 37-342, 37-395, 38-270, 38-275, 38-280, 38-297, 38-353, 39-271, 39-313, 42-280, 42-294, 43-362, 45-134, 45-288, 45-310, 45-313, 45-314, 45-360, 46-164, 46-251, 46-278, 46-299, 46-300, 46-303, 46-306, 46-309, 46-310, 46-314, 46-319, 47-265, 47-287, 47-340, 47-371, 49-304, 50-269, 50-301, 50-337, 50-395, 51-234, 51-271, 51-280, 51-309, 51-311, 51-345, 51-680, 52-270, 52-274, 52-276, 52-322, 53-271, 53-278, 53-322, 53-327, 54-287, 55-347, 56-326, 57-301, 57-325, 57-327, 57-343, 60-223, 61-327, 61-355, 61-395, 62-330, 62-349, 63-298, 63-327, 63-344, 65-293, 65-317, 65-324, 68-340,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	70-350, 77-212, 85-316, 85-332, 103-378, 108-395, 111-386, 124-420, 136-371, 137-392, 154-382, 337-610, 337-616, 396-1013, 396-1022, 396-1026, 396-1039, 398-987, 399-1022, 400-1014, 402-1057, 408-1012, 412-1039, 413-1039, 421-852, 423-1039, 430-808, 436-1050, 439-813, 446-1039, 447-997, 447-1046, 453-925, 454-1041, 457-1037, 474-1003, 476-793, 477-1052, 479-1046, 490-863, 490-1013, 493-738, 498-748, 498-1046, 499-751, 499-766, 501-1046, 502-1046, 503-970, 503-1034, 504-906, 505-1019, 506-1078, 514-1052, 520-989, 520-1050, 527-792, 528-1044, 531-1046, 542-784, 544-1039, 545-811, 548-737, 548-754, 548-763, 563-1038, 563-1052, 564-1043, 565-1040, 567-1052, 569-1041, 571-1046, 573-1035, 575-1041, 576-1036, 577-1046, 577-1049, 578-1040, 579-1052, 580-1040, 581-1045, 582-1042, 583-792, 585-1039, 586-809, 586-880, 587-1035, 587-1049, 589-1040, 589-1041, 590-1043, 591-1034, 593-1042, 594-1040, 594-1046, 595-1034, 598-902, 598-1043, 599-1042, 599-1061, 599-1062, 601-1034,
	602-1035, 609-940, 612-1036, 612-1049, 612-1064, 614-1039, 616-1046, 619-1040, 623-1040, 623-1042, 623-1066, 624-1048, 625-1040, 626-1040, 627-1040, 627-1043, 628-1035, 628-1040, 630-1035, 634-1034, 635-1034, 635-1086, 636-1045, 636-1052, 639-1046, 642-1040, 644-1052, 652-1040, 655-1043, 657-1034, 658-917, 658-1020, 662-1043, 664-807, 665-1062, 667-1035, 668-1035, 681-1050, 687-1045, 688-1039, 695-1037, 696-1040, 697-1034, 716-970, 724-1034, 730-1040, 731-1042, 732-1029, 732-1040, 734-906, 735-1010, 735-1044, 738-1033, 740-1040, 758-1040, 773-1040, 784-1039, 791-1037, 796-1040, 804-1036, 814-1055, 815-1046, 830-1035, 843-1040, 869-1044
99/ 7510361CB1/ 1099	1-255, 1-256, 1-386, 1-445, 1-453, 1-469, 1-534, 1-552, 1-555, 1-563, 1-570, 1-572, 1-586, 1-598, 1-609, 1-610, 1-891, 3-192, 49-457, 57-449, 57-458, 72-185, 83-575, 204-421, 215-407, 233-755, 233-797, 268-342, 268-469, 280-458, 289-489, 356-592, 388-489, 413-545, 425-488, 470-610, 472-1099, 573-610
100/ 7507013CB1/ 1110	1-137, 1-223, 1-245, 1-1110, 5-224, 11-245, 115-587, 265-379, 481-688, 481-1108, 491-902, 491-1042, 497-952, 498-732, 498-1017, 519-728, 522-656, 535-939, 546-1073, 563-1110, 566-1045, 568-1088, 575-1067, 582-1092, 624-764, 638-1108, 649-1107, 650-1107, 658-1107, 659-1108, 665-889, 673-1105, 682-748, 712-1045, 714-1110, 726-1110, 738-1107, 746-1107, 748-1107, 782-1107, 793-1107, 799-1076, 843-1057, 845-1107, 868-1069, 924-1108

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
101/ 7510507CB1/ 3082	1-313, 1-407, 1-431, 1-434, 1-450, 1-516, 1-580, 1-3075, 1-3080, 156-458, 156-675, 158-723, 173-723, 178-384, 186-847, 388-900, 389-901, 527-983, 559-1283, 568-1237, 593-1214, 636-913, 672-935, 695-1284, 728-1031, 743-1147, 749-1334, 763-1156, 796-1359, 898-1290, 898-1482, 898-1507, 898-1555, 918-1423, 960-1276, 973-1274, 1016-1249, 1316-1817, 1316-2030, 1324-2046, 1326-2075, 1352-1631, 1352-1944, 1352-2030, 1381-2124, 1384-1974, 1404-1764, 1439-2234, 1440-2119, 1454-1974, 1476-1980, 1490-1658, 1513-1848, 1550-2277, 1561-2097, 1586-2127, 1615-1998, 1644-2062, 1657-2171, 1762-2129, 1764-2087, 1816-2090, 1819-2058, 1819-2275, 1831-2482, 1889-2141, 1917-2168, 1936-2199, 1936-2201, 1941-2200, 1952-2330, 1958-2589, 1964-2230, 1991-2499, 2008-2231, 2038-2264, 2063-2465, 2063-2468, 2090-2326, 2111-2400, 2115-2679, 2132-2674, 2149-2432, 2157-2390, 2158-2340, 2167-2410, 2245-2912, 2256-2708, 2270-2494, 2270-2723, 2299-2636, 2317-3080, 2337-2717,
	2340-2562, 2340-2580, 2340-2860, 2340-2951, 2340-3032, 2352-3001, 2360-3063, 2365-3050, 2375-2882, 2379-3060, 2380-2624, 2397-2653, 2428-2708, 2429-2680, 2449-3047, 2451-3047, 2451-3064, 2463-3080, 2473-3024, 2476-2714, 2493-3080, 2505-3067, 2548-3080, 2563-3080, 2573-2718, 2573-2817, 2588-2949, 2613-2978, 2619-3067, 2620-3065, 2621-3065, 2632-3080, 2640-3069, 2643-3065, 2649-2909, 2652-3070, 2653-3065, 2659-3064, 2666-3065, 2686-3069, 2721-2956, 2721-3076, 2729-3069, 2734-2847, 2753-3063, 2762-3068, 2763-3067, 2773-3063, 2780-3052, 2782-3069, 2797-3039, 2799-3057, 2804-3082, 2807-3082, 2813-3072, 2820-2970, 2820-3069, 2850-3073, 2861-3051, 2867-3066, 2910-3080
102/ 90106370CB1/ 1548	1-580, 1-601, 1-648, 1-743, 1-747, 1-764, 1-771, 1-774, 1-779, 1-781, 1-822, 1-825, 1-844, 1-846, 1-852, 1-854, 1-857, 1-869, 2-579, 2-700, 2-778, 5-808, 5-869, 9-869, 31-869, 36-869, 37-869, 43-516, 43-604, 43-680, 43-685, 43-705, 43-711, 43-735, 43-803, 43-834, 43-852, 43-854, 71-869, 73-869, 88-869, 159-869, 166-869, 231-869, 234-987, 265-869, 351-577, 627-1367, 634-1269, 662-1532, 712-1424, 751-862, 766-1239, 781-1548

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
52	7506904CB1	CORPNOT02
53	7506909CB1	KIDNNOT05
54	7507096CB1	KIDNNOC01
55	7507098CB1	KIDNNOC01
56	7507099CB1	KIDNNOC01
57	7501399CB1	BRACDIK08
58	7504768CB1	PROSDIT01
59	7500757CB1	BRAINOT09
60	1730616CB1	FIBPFEN06
61	190404CB1	SYNORAB01
62	7500679CB1	LUNLTMT01
63	7500687CB1	BRAENOT02
64	7500688CB1	BRAENOT02
65	7500697CB1	KIDNNOT32
66	7500709CB1	SINTNOR01
67	7500711CB1	SINTNOR01
68	7500723CB1	EOSITXT01
69	7500764CB1	PENITUT01
70	7500772CB1	SINTNOR01
71	7501350CB1	MONOTXT02
72	7506396CB1	LUNGAST01
73	7505917CB1	SINTFER02
74	7500701CB1	BRAPDIT01
75	7500702CB1	BRSTNOT09
76	6044343CB1	BRAUTDR02
77	7503990CB1	BRSTNOT05
78	7504655CB1	PANCNOT01
79	7504690CB1	NERDTDN03
80	7504720CB1	NGANNOT01
81	7504722CB1	293TF3T01
82	7504733CB1	COLNFET02
83	7507100CB1	KIDNNOC01
84	7503330CB1	KIDNFEE02
85	7504519CB1	LNODNOT05
86	7504705CB1	MYOMNOT01
87	7504738CB1	MPHGLPT02
88	7510280CB1	BRABNOE02
89	7503700CB1	LUNGNOT02
90	7504685CB1	ADRENOT07
91	7506844CB1	LUNGFET05
92	7510259CB1	null lib
93	7510444CB1	LIVRTMR01
94	7510494CB1	MPHGNOT03
95	6486485CB1	MIXDUNB01
96	7503772CB1	PROSNOT28
97	7503773CB1	ENDCNOT03
98	7504698CB1	HEARNON08
99	7510361CB1	LSUBNOT03
100	7507013CB1	BRSTNOT02

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
101	7510507CB1	SCOMDIC01
102	90106370CB1	SPLNDIC01

Table 6

Library	Vector	Library Description
293TF3T01	pINCY	Library was constructed using RNA isolated from a serum-starved transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue. The cells were transformed with adenovirus 5 DNA.
ADRENOT07	pINCY	Library was constructed using RNA isolated from adrenal tissue removed from a 61-year-old female during a bilateral adrenalectomy. Patient history included an unspecified disorder of the adrenal glands.
BRABNOE02	PBK-CMV	This 5' biased random primed library was constructed using RNA isolated from vermis tissue removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly, and an enlarged spleen and liver. Patient medications included simethicone, Lasix, Digoxin, Colace, Zantac, captopril, and Vasotec.
BRACDIK08	PSPORT1	This amplified and normalized library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 57-year-old Caucasian male who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day for 40 years).
BRAENOT02	pINCY	Library was constructed using RNA isolated from posterior parietal cortex tissue removed from the brain of a 35-year-old Caucasian male who died from cardiac failure.
BRAINOT09	pINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.
BRAPDIT01	pINCY	Library was constructed using RNA isolated from diseased pons tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Serology was negative. Patient history included Huntington's disease, emphysema, and tobacco abuse.
BRAUTDR02	PCDNA2.1	This random primed library was constructed using RNA isolated from pooled amygdala and entorhinal cortex tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRSTNOT02	PSPORT1	Library was constructed using RNA isolated from diseased breast tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated proliferative fibrocysytic changes characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Pathology for the associated tumor tissue indicated an invasive grade 4 mammary adenocarcinoma. Patient history included atrial tachycardia and a benign neoplasm. Family history included cardiovascular and cerebrovascular disease.

Table 6

Library	Vector	Library Description
BRSTNOT05	PSPORT1	Library was constructed using RNA isolated from breast tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated multicentric invasive grade 4 lobular carcinoma. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular and cardiovascular disease, breast and prostate cancer, and type I diabetes.
BRSTNOT09	pINCY	Library was constructed using RNA isolated from breast tissue removed from a 45-year-old Caucasian female during unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated invasive nuclear grade 2-3 adenocarcinoma, with 3 of 23 lymph nodes positive for metastatic disease. Immunostains for estrogen/progesterone receptors were positive, and uninvolved tissue showed proliferative changes. The patient concurrently underwent a total abdominal hysterectomy. Patient history included valvuloplasty of mitral valve without replacement, rheumatic mitral insufficiency, and rheumatic heart disease. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
COLNFET02	pINCY	Library was constructed using RNA isolated from the colon tissue of a Caucasian female fetus, who died at 20 weeks' gestation.
CORPNOT02	pINCY	Library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.
ENDCNOT03	pINCY	Library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a neonatal Caucasian male.
EOSITXT01	pINCY	Library was constructed using RNA isolated from eosinophils stimulated with IL-5.
FIBPFEN06	pINCY	The normalized prostate stromal fibroblast tissue libraries were constructed from 1.56 million independent clones from a prostate fibroblast library. Starting RNA was made from fibroblasts of prostate stroma removed from a male fetus, who died after 26 weeks' gestation. The libraries were normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48-hours/round)reannealing hybridization was used. The library was then linearized and recircularized to select for insert containing clones as follows: plasmid DNA was prepped from approximately 1 million clones from the normalized prostate stromal fibroblast tissue libraries following soft agar transformation.
HEARNON08	PBLUESCRIPT	This normalized, pooled mixed untreated cardiomyocytes and heart tissue library was constructed from 1 million independent clones from a mixed heart and cardiomyocyte library. Starting RNA was made from polyA RNA isolated from pooled untreated cardiomyocytes removed from a 16-week-old Caucasian fetus (donor A) and heart tissue removed from a 21-year-old Caucasian female (donor B) who died from cardiopulmonary arrest. Donor B's history included delivery four months prior to death and seizures. Patient medications included unspecified birth control pills. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

Table 6

Library	Vector	Library Description
KIDNFEE02	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from kidney tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation. Serology was negative.
KIDNNOC01	pINCY	This large size-fractionated library was constructed using RNA isolated from pooled left and right kidney tissue removed from a Caucasian male fetus, who died from Patau's syndrome (trisomy 13) at 20-weeks' gestation.
KIDNNOT05	PSPORT1	Library was constructed using RNA isolated from the kidney tissue of a 2-day-old Hispanic female, who died from cerebral anoxia. Family history included congenital heart disease.
KIDNNOT32	pINCY	Library was constructed using RNA isolated from kidney tissue removed from a 49-year-old Caucasian male who died from an intracranial hemorrhage and cerebrovascular accident. Patient history included tobacco abuse.
LIVRTMR01	PCDNA2.1	This random primed library was constructed using RNA isolated from liver tissue removed from a 62-year-old Caucasian female during partial hepatectomy and exploratory laparotomy. Pathology for the matched tumor tissue indicated metastatic intermediate grade neuroendocrine carcinoma, consistent with islet cell tumor, forming nodules ranging in size, in the lateral and medial left liver lobe. The pancreas showed fibrosis, chronic inflammation and fat necrosis consistent with pseudocyst. The gallbladder showed mild chronic cholecystitis. Patient history included malignant neoplasm of the pancreas tail, pulmonary embolism, hyperlipidemia, thrombophlebitis, joint pain in multiple joints, type II diabetes, benign hypertension, cerebrovascular disease, and normal delivery. Previous surgeries included distal pancreatectomy, total splenectomy, and partial hepatectomy. Family history included pancreas cancer with secondary liver cancer, benign hypertension, and hyperlipidemia.
LNODNOT05	pINCY	Library was constructed using RNA isolated from lymph node tissue obtained from a 14-year-old Caucasian female, who died from cardiac arrest secondary to burns. Serology was negative.
LSUBNOT03	pINCY	Library was constructed using RNA isolated from submandibular gland tissue obtained from a 68-year-old Caucasian male during a sialoadenectomy. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
LUNGAST01	PSPORT1	Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male, who died from head trauma. Patient history included asthma.
LUNGFET05	PSPORT1	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation from anencephalus.
LUNGNOT02	PBLUESCRIPT	Library was constructed using RNA isolated from the lung tissue of a 47-year-old Caucasian male, who died of a subarachnoid hemorrhage.

Table 6

Library	Vector	Library Description
LUNLTMT01	pINCY	The library was constructed using RNA isolated from right middle lobe lung tissue removed from a 63-year-old Caucasian female during a segmental lung resection. Pathology for the associated tumor tissue indicated grade3 adenocarcinoma in the right lower lobe and right middle lobe that infiltrated the parietal pleural surface. Metastatic grade 3 adenocarcinoma was found in the diaphragm. The lymph nodes contained metastatic grade 3 adenocarcinoma and involved the superior mediastinal and inferior mediastinal lymph nodes. Patient history included hyperlipidemia. Family history included benign hypertension,cerebrovascular disease, breast cancer, and hyperlipidemia.
MIXDUNB01	pINCY	Library was constructed using RNA isolated from myometrium removed from a 41-year-old Caucasian female (A) during vaginal hysterectomy with a dilatation and curettage and untreated smooth muscle cells removed from the renal vein of a 57-year-old Caucasian male. Pathology for donor A indicated the myometrium and cervix were unremarkable. The endometrium was secretory and contained fragments of endometrial polyps. Benign endo- and ectocervical mucosa were identified in the endocervix. Pathology for the associated tumor tissue indicated uterine leiomyoma. Medical history included an unspecified menstrual disorder, ventral hernia, normal delivery, a benign ovarian neoplasm, and tobacco abuse in donor A. Previous surgeries included a bilateral destruction of fallopian tubes, removal of a solitary ovary, and an exploratory laparotomy in donor A. Medications included ferrous sulfate in donor A.
MONOTXT02	pINCY	The library was constructed using RNA isolated from treated monocytes from peripheral blood removed from a 42-year-old female. The cells were treated with interleukin-10 (IL-10) and lipopolysaccharide (LPS). IL-10 was added at time 0 at 10 ng/ml, LPS was added at 1 hour at 5 ng/ml. The monocytes were isolated from buffy coat by adherence to plastic. Incubation time was 24 hours.
MPHGLPT02	PSPORT1	Library was constructed using RNA isolated from adherent mononuclear cells, which came from a pool of male and female donors. The cells were stimulated with LPS.
MPHGNOT03	PBLUESCRIPT	Library was constructed using RNA isolated from plastic adherent mononuclear cells isolated from buffy coat units obtained from unrelated male and female donors.
MYOMNOT01	PSPORT1	Library was constructed using RNA isolated from uterine myometrial tissue removed from a 43-year-old Caucasian female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. Family history included lung cancer, stroke, type II diabetes, hepatic lesion, chronic liver disease, hyperlipidemia, congenital heart anomaly, and mitral valve prolapse.

Table 6

Library	Vector	Library Description
NERDTDN03	pINCY	This normalized dorsal root ganglion tissue library was constructed from 1.05 million independent clones from a dorsal root ganglion tissue library. Starting RNA was made from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema, acute bronchopneumonia, bilateral pleural effusions, pericardial effusion, and malignant lymphoma (natural killer cell type). The patient presented with pyrexia of unknown origin, malaise, fatigue, and gastrointestinal bleeding. Patient history included probable cytomegalovirus infection, liver congestion, and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, respiratory failure, pneumonia of the left lung, natural killer cell lymphoma of the pharynx, Bell's palsy, and tobacco and alcohol abuse. Previous surgeries included colonoscopy, closed colon biopsy, adenotonsillectomy, and nasopharyngeal endoscopy and biopsy. Patient medications included Diflucan (fluconazole), Deltasone (prednisone), hydrocodone, Lortab, Alprazolam, Reazodone, ProMace-Cytabom, Etoposide, Cisplatin, Cytarabine, and dexamethasone. The patient received radiation therapy and multiple blood transfusions. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
NGANNOT01	PSPORT1	Library was constructed using RNA isolated from tumorous neuroganglion tissue removed from a 9-year-old Caucasian male during a soft tissue excision of the chest wall. Pathology indicated a ganglioneuroma. Family history included asthma.
PANCNOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the pancreatic tissue of a 29-year-old Caucasian male who died from head trauma.
PENITUT01	pINCY	Library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Patient history included benign neoplasm of the large bowel, atherosclerotic coronary artery disease, angina pectoris, gout, and obesity. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.
PROSDIT01	pINCY	The library was constructed using RNA isolated from diseased prostate tissue removed from a 58-year-old Caucasian male during radical prostatectomy, regional lymph node excision, and prostate needle biopsy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma Gleason grade 3+3, which formed a predominant mass involving the right posterior superior prostate. Another microscopic focus of tumor was identified in the left posterior inferior. The tumor invaded the capsule but didnot extend beyond it. The patient presented with elevated prostate specific antigen (PSA), nocturia, hematuria, and induration. Patient history included benign hypertension. Family history included benign hypertension and prostate cancer.

Table 6

Library	Vector	Library Description
PROSNOT28	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 55-year-old Caucasian male during a radical prostatectomy and regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma, Gleason grade 5+4. The patient presented with elevated prostate specific antigen (PSA). Family history included lung and breast cancer.
SCOMDIC01	PSPORT1	This large size-fractionated library was constructed using RNA isolated from diseased spinal cord tissue removed from the base of the medulla of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day, for 40 years).
SINTFER02	pINCY	This random primed library was constructed using RNA isolated from small intestine tissue removed from a Caucasian male fetus who died from fetal demise.
SINTNOR01	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 31-year-old Caucasian female during Roux-en-Y gastric bypass. Patient history included clinical obesity.
SPLNDIC01	pINCY	This large size-fractionated library was constructed using pooled cDNA from two different donors. cDNA was generated using mRNA isolated from spleen tissue removed from an 8-year-old Black male (donor A) who died from anoxia and from diseased spleen tissue removed from a 14-year-old Asian male (donor B) during a total splenectomy. Pathology for donor B indicated changes consistent with idiopathic thrombocytopenic purpura. Serologies were negative for donor A. Donor B presented with bruising. Patient medications included DDAVP, Versed, labetalol (donor A), and Vincristine (donor B).
SYNORAB01	PBLUESCRIPT	Library was constructed using RNA isolated from the synovial membrane tissue of a 68-year-old Caucasian female with rheumatoid arthritis.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value = 1.0E-8 or less; Full Length sequences: Probability value = 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value = 1.06E-6; Assembled ESTs: fasta Identity = 95% or greater and Match length = 200 bases or greater; fastx E value = 1.0E-8 or less; Full Length sequences: fastx score = 100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-	Probability value = 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM, INCY, SMART and TIGRFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-	PFAM, INCY, SMART or TIGRFAM hits: Probability value = 1.0E-3 or less; Signal peptide hits: Score = 0 or greater
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score ≥ GCG specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.

Table 7

Program	Description	Reference	Parameter Threshold
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score = 3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. On Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence (AAAI) Press, Menlo Park, CA, and MIT Press, Cambridge, MA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
52	7506904	1700625H1	SNP00039444	77	1249	A	A	G	noncoding	n/d	n/a	n/a	n/a
52	7506904	1700625H1	SNP00072668	152	1324	A	A	G	noncoding	n/d	n/a	n/a	n/a
52	7506904	6331607H1	SNP00039444	138	1239	A	A	G	noncoding	n/d	n/a	n/a	n/a
52	7506904	6331607H1	SNP00072668	213	1314	G	A	G	noncoding	n/d	n/a	n/a	n/a
53	7506909	1700625H1	SNP00039444	77	947	A	A	G	noncoding	n/d	n/a	n/a	n/a
53	7506909	1700625H1	SNP00072668	152	1022	A	A	G	noncoding	n/d	n/a	n/a	n/a
53	7506909	6331607H1	SNP00039444	138	937	A	A	G	noncoding	n/d	n/a	n/a	n/a
53	7506909	6331607H1	SNP00072668	213	1012	G	A	G	noncoding	n/d	n/a	n/a	n/a
54	7507096	1560345H1	SNP00047855	58	1755	T	T	C	noncoding	0.59	0.61	0.54	0.69
54	7507096	3213196H1	SNP00129679	16	1008	T	C	T	noncoding	n/a	n/a	n/a	n/a
55	7507098	1560345H1	SNP00047855	58	1900	T	T	C	noncoding	0.59	0.61	0.54	0.69
55	7507098	3213196H1	SNP00129679	16	1153	T	C	T	noncoding	n/a	n/a	n/a	n/a
56	7507099	1560345H1	SNP00047855	58	1677	T	T	C	noncoding	0.59	0.61	0.54	0.69
56	7507099	3213196H1	SNP00129679	16	930	T	C	T	noncoding	n/a	n/a	n/a	n/a
57	7501399	1273836H1	SNP00100328	155	1137	G	G	A	noncoding	n/a	n/a	n/a	n/a
57	7501399	3744343H1	SNP00100327	9	18	C	C	T	noncoding	n/a	n/a	n/a	n/a
57	7501399	3744344H1	SNP00100327	7	16	C	C	T	noncoding	n/a	n/a	n/a	n/a
57	7501399	4784901H2	SNP00068331	206	1687	C	C	T	noncoding	n/a	n/a	n/a	n/a
57	7501399	5353388H1	SNP00068331	118	1688	C	C	T	noncoding	n/a	n/a	n/a	n/a
57	7501399	6404778H1	SNP00068330	29	1310	C	T	C	noncoding	0.12	n/a	n/a	n/a
57	7501399	6869706H1	SNP00128804	604	1197	G	G	T	noncoding	n/a	n/a	n/a	n/a
57	7501399	6869706H1	SNP00068330	488	1313	T	T	C	noncoding	0.12	n/a	n/a	n/a
58	7504768	1343743H1	SNP00052446	160	391	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1343743H1	SNP00067987	200	431	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1343743H1	SNP00068446	107	338	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	1343743H1	SNP00075667	114	345	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1343743H1	SNP00132838	56	287	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1343757H1	SNP00068474	100	644	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1343782H1	SNP00063753	28	24	C	C	A	P4	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	1344211H1	SNP00044426	20	224	G	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	1344444H1	SNP00056531	187	631	C	C	A	noncoding	n/d	n/a	n/a	n/a
58	7504768	1344448H1	SNP00054416	20	46	T	T	C	F11	n/a	n/a	n/a	n/a
58	7504768	1345066H1	SNP00146179	63	64	C	C	A	P17	n/a	n/a	n/a	n/a
58	7504768	1345133H1	SNP00129147	240	702	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1345518H1	SNP00063754	51	185	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	1345518H1	SNP00140832	78	212	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1345643H1	SNP00063754	96	181	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	1345643H1	SNP00076561	28	113	G	G	C	P33	n/a	n/a	n/a	n/a
58	7504768	1345643H1	SNP00132838	198	283	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1345643H1	SNP00140832	123	208	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1345659H1	SNP00047368	59	61	T	C	T	V16	n/a	n/a	n/a	n/a
58	7504768	1346287H1	SNP00054416	45	47	T	T	C	S11	n/a	n/a	n/a	n/a
58	7504768	1346475H1	SNP00037414	77	528	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1347022H1	SNP00037414	54	529	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1347466H1	SNP00054416	43	45	T	T	C	S11	n/a	n/a	n/a	n/a
58	7504768	1347763H1	SNP00076561	23	117	G	G	C	A35	n/a	n/a	n/a	n/a
58	7504768	1348038H1	SNP00011894	66	721	T	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1348543H1	SNP00146294	175	280	C	C	G	noncoding	n/a	n/a	n/a	n/a
58	7504768	1492269H1	SNP00054365	89	562	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	1493665H1	SNP00063753	36	35	C	C	A	F7	n/a	n/a	n/a	n/a
58	7504768	1495160H1	SNP00009321	64	456	C	T	C	noncoding	0.92	n/a	n/a	n/a
58	7504768	1495160H1	SNP00037413	35	427	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1542723H1	SNP00047368	66	55	C	C	T	S14	n/a	n/a	n/a	n/a
58	7504768	1543416H1	SNP00011894	183	722	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1543833H1	SNP00043769	21	607	G	G	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1544755H1	SNP00054365	62	563	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	1546401H1	SNP00054365	14	564	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	1549314H1	SNP00068446	172	348	G	G	A	noncoding	0.99	n/d	n/d	n/d

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	1549314H1	SNP00075667	179	355	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1549314H1	SNP00132838	121	297	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1549314H1	SNP00140832	46	222	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1549323H1	SNP00011894	182	718	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1549323H1	SNP00129147	163	699	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1549587H1	SNP00063754	92	194	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	1549587H1	SNP00076561	24	126	G	G	C	A38	n/a	n/a	n/a	n/a
58	7504768	1549587H1	SNP00140832	119	221	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1549671H1	SNP00056531	196	630	C	C	A	noncoding	n/d	n/a	n/a	n/a
58	7504768	1682001H1	SNP00042307	30	314	C	C	A	noncoding	0.20	0.07	0.08	0.41
58	7504768	1682982H1	SNP00037414	29	530	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1683569H1	SNP00076561	115	116	G	G	C	T34	n/a	n/a	n/a	n/a
58	7504768	1683572H1	SNP00047368	56	63	C	C	T	Q17	n/a	n/a	n/a	n/a
58	7504768	1683572H1	SNP00046952	63	72	G	A	G	A20	n/a	n/a	n/a	n/a
58	7504768	1685714H1	SNP00044426	103	244	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	1686274H1	SNP00046952	75	71	G	A	G	A19	n/a	n/a	n/a	n/a
58	7504768	1687168H1	SNP00054365	25	561	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	1688106H1	SNP00052446	77	410	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1688106H1	SNP00067987	117	449	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1688106H1	SNP00068446	24	357	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	1688106H1	SNP00075667	31	364	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1688329H1	SNP00011894	100	719	T	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1690353H1	SNP00155250	68	373	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1690881H1	SNP00044426	226	225	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	1709318H1	SNP00056531	188	632	C	C	A	noncoding	n/d	n/a	n/a	n/a
58	7504768	1711676H1	SNP00046952	73	73	G	A	G	G20	n/a	n/a	n/a	n/a
58	7504768	1724316H1	SNP00044426	54	226	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	1725201H1	SNP00054416	70	70	T	T	C	V19	n/a	n/a	n/a	n/a
58	7504768	1808532H1	SNP00042307	179	185	C	C	A	noncoding	0.20	0.07	0.08	0.41

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	1809090H1	SNP00046952	76	75	G	A	G	V21	n/a	n/a	n/a	n/a
58	7504768	1809090H1	SNP00054416	53	49	T	T	C	V12	n/a	n/a	n/a	n/a
58	7504768	1809123H1	SNP00044426	92	234	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	1809992H1	SNP00052446	115	390	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1809992H1	SNP00067987	155	430	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1810488H1	SNP00052446	86	386	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1810488H1	SNP00067987	126	426	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1810488H1	SNP00068446	33	333	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	1810488H1	SNP00075667	40	340	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1811174H1	SNP00047368	70	60	C	C	T	R16	n/a	n/a	n/a	n/a
58	7504768	1814827H1	SNP00044426	272	230	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	1855821H1	SNP00129147	172	700	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1856907H1	SNP00046892	39	23	T	T	C	V3	0.06	0.04	0.04	0.03
58	7504768	1856907H1	SNP00124735	72	56	G	G	A	W14	n/a	n/a	n/a	n/a
58	7504768	1857613H1	SNP00046892	17	14	T	T	C	noncoding	0.06	0.04	0.04	0.03
58	7504768	1857613H1	SNP00124735	50	47	G	G	A	S11	n/a	n/a	n/a	n/a
58	7504768	1913348H1	SNP00044426	152	228	G	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	1913858H1	SNP00063753	35	25	C	C	A	P4	n/a	n/a	n/a	n/a
58	7504768	1914285H1	SNP00052446	68	389	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1914285H1	SNP00068446	15	336	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	1914285H1	SNP00075667	22	343	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1915830H1	SNP00046952	82	80	G	A	G	L22	n/a	n/a	n/a	n/a
58	7504768	1915830H1	SNP00054416	59	54	T	T	C	W14	n/a	n/a	n/a	n/a
58	7504768	1917204H1	SNP00052446	179	400	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	1917204H1	SNP00067987	219	440	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1917204H1	SNP00068446	126	347	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	1917204H1	SNP00075667	133	354	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1917204H1	SNP00132838	75	296	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	1917215H1	SNP00046952	91	90	G	A	G	A26	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	1917215H1	SNP00054416	68	65	T	T	C	H17	n/a	n/a	n/a	n/a
58	7504768	2068818H1	SNP00054365	36	586	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	2187806H1	SNP00011894	91	720	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	2187806H1	SNP00129147	72	701	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	2233404H1	SNP00068474	36	643	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	2271117H1	SNP00063753	35	34	C	C	A	S7	n/a	n/a	n/a	n/a
58	7504768	2276979H1	SNP00052446	29	415	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	2276979H1	SNP00067987	69	455	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	2388554H1	SNP00063753	37	36	C	C	A	L8	n/a	n/a	n/a	n/a
58	7504768	2941522H1	SNP00056531	188	653	C	C	A	noncoding	n/d	n/a	n/a	n/a
58	7504768	3272323H1	SNP00063753	25	22	C	C	A	A3	n/a	n/a	n/a	n/a
58	7504768	3272538H1	SNP00063753	33	23	C	C	A	V3	n/a	n/a	n/a	n/a
58	7504768	3272748H1	SNP00063753	24	21	C	C	A	L3	n/a	n/a	n/a	n/a
58	7504768	3272754H1	SNP00063753	17	14	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3272891H1	SNP00047368	35	59	T	C	T	I15	n/a	n/a	n/a	n/a
58	7504768	3273142H1	SNP00068474	33	641	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3273256H1	SNP00052446	161	387	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3273256H1	SNP00067987	201	427	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3273256H1	SNP00068446	108	334	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	3273256H1	SNP00075667	115	341	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3273256H1	SNP00132838	57	284	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3276649H1	SNP00063754	181	182	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	3276649H1	SNP00076561	113	114	G	G	C	A34	n/a	n/a	n/a	n/a
58	7504768	3276649H1	SNP00140832	208	209	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3276751H1	SNP00067987	163	429	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3276751H1	SNP00132838	19	285	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3276756H1	SNP00063753	18	16	C	C	A	T1	n/a	n/a	n/a	n/a
58	7504768	3277293H1	SNP00052446	25	382	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3277293H1	SNP00067987	65	422	C	C	T	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	3316682H1	SNP00009321	44	455	C	T	C	noncoding	0.92	n/a	n/a	n/a
58	7504768	3316682H1	SNP00037413	15	426	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3317420H1	SNP00068446	123	337	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	3317420H1	SNP00075667	130	344	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3317420H1	SNP00132838	72	286	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3317786H1	SNP00063753	23	20	C	C	A	C2	n/a	n/a	n/a	n/a
58	7504768	3318147H1	SNP00009321	45	452	T	T	C	noncoding	0.92	n/a	n/a	n/a
58	7504768	3318147H1	SNP00037413	16	423	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3319744H1	SNP00063833	140	133	G	G	C	G40	n/d	n/d	n/d	n/d
58	7504768	3351152H1	SNP00054416	41	44	T	T	C	L10	n/a	n/a	n/a	n/a
58	7504768	3356872H1	SNP00063754	54	184	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	3356872H1	SNP00140832	81	211	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3358153H1	SNP00063753	23	19	C	C	A	S2	n/a	n/a	n/a	n/a
58	7504768	3358196H1	SNP00047368	61	57	T	C	T	F15	n/a	n/a	n/a	n/a
58	7504768	3359959H1	SNP00068446	270	339	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	3360072H1	SNP00011894	86	715	T	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3360072H1	SNP00129147	67	696	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3361360H1	SNP00046952	82	70	G	A	G	G19	n/a	n/a	n/a	n/a
58	7504768	3362266H1	SNP00046952	80	76	G	A	G	G21	n/a	n/a	n/a	n/a
58	7504768	3362266H1	SNP00054416	57	50	T	T	C	V12	n/a	n/a	n/a	n/a
58	7504768	3362413H1	SNP00047368	75	65	C	C	T	H17	n/a	n/a	n/a	n/a
58	7504768	3362834H1	SNP00037414	46	559	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3362834H1	SNP00054365	80	593	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	3365106H1	SNP00140832	194	210	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3497647H1	SNP00068474	85	642	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3498973H1	SNP00054416	70	59	T	T	C	I15	n/a	n/a	n/a	n/a
58	7504768	3499611H1	SNP00037414	71	524	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3962030H1	SNP00146294	137	277	C	C	G	noncoding	n/a	n/a	n/a	n/a
58	7504768	3962418H1	SNP00146294	17	279	C	C	G	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	3962841H1	SNP00063754	184	183	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	3962841H1	SNP00076561	116	115	G	G	C	S34	n/a	n/a	n/a	n/a
58	7504768	3963802H1	SNP00132838	58	236	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3964102H1	SNP00052446	74	388	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3964102H1	SNP00067987	114	428	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3964673H1	SNP00042307	39	312	C	C	A	noncoding	0.20	0.07	0.08	0.41
58	7504768	3964682H1	SNP00054416	61	58	T	T	C	I15	n/a	n/a	n/a	n/a
58	7504768	3965811H1	SNP00011894	82	716	T	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3965811H1	SNP00129147	63	697	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3966081H1	SNP00076561	50	112	G	G	C	R33	n/a	n/a	n/a	n/a
58	7504768	3966329H1	SNP00146294	131	278	C	C	G	noncoding	n/a	n/a	n/a	n/a
58	7504768	3966939H1	SNP00068446	62	335	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	3967020H1	SNP00044426	104	235	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	3967122H1	SNP00037414	53	527	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3967837H1	SNP00155250	173	371	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3968413H1	SNP00011894	210	726	T	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3968413H1	SNP00129147	191	707	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3968755H1	SNP00037414	95	612	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3968755H1	SNP00054365	129	646	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	3978889H1	SNP00052446	90	361	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	3978889H1	SNP00067987	130	401	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	3978889H1	SNP00068446	37	309	G	G	A	noncoding	0.99	n/d	n/d	n/d
58	7504768	4108724H1	SNP00043769	18	606	G	G	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4109074H1	SNP00009321	46	454	T	T	C	noncoding	0.92	n/a	n/a	n/a
58	7504768	4109695H1	SNP00043769	5	602	G	G	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4110760H1	SNP00054416	60	52	T	T	C	M13	n/a	n/a	n/a	n/a
58	7504768	4111002H1	SNP00124735	55	51	G	G	A	A13	n/a	n/a	n/a	n/a
58	7504768	4111048H1	SNP00054365	45	583	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	4111426H1	SNP00075667	5	342	C	C	T	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	4111558H1	SNP00047368	60	56	C	C	T	C14	n/a	n/a	n/a	n/a
58	7504768	4111717H1	SNP00009321	9	453	T	T	C	noncoding	0.92	n/a	n/a	n/a
58	7504768	4112155H1	SNP00155250	200	370	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	4112893H1	SNP00047368	66	54	C	C	T	R14	n/a	n/a	n/a	n/a
58	7504768	4113023H1	SNP00054416	83	79	T	T	C	F22	n/a	n/a	n/a	n/a
58	7504768	4276507H1	SNP00155250	93	372	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	4278118H1	SNP00042307	34	310	C	C	A	noncoding	0.20	0.07	0.08	0.41
58	7504768	4404036H1	SNP00054416	59	51	T	T	C	S13	n/a	n/a	n/a	n/a
58	7504768	4404477H1	SNP00047368	61	58	T	C	T	I15	n/a	n/a	n/a	n/a
58	7504768	4404731H1	SNP00075667	157	346	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4405580H1	SNP00037414	140	535	T	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4405580H1	SNP00054365	174	569	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	4408070H1	SNP00046952	78	94	G	A	G	S27	n/a	n/a	n/a	n/a
58	7504768	4569531H1	SNP00068474	28	640	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4569542H1	SNP00044426	81	229	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	4574516H1	SNP00076561	183	137	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	4593560H1	SNP00063833	9	134	G	G	C	A40	n/d	n/d	n/d	n/d
58	7504768	4595478H1	SNP00037414	88	533	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4595478H1	SNP00054365	122	567	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	4641606H1	SNP00067987	230	432	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4642115H1	SNP00046892	25	22	C	T	C	A3	0.06	0.04	0.04	0.03
58	7504768	4642115H1	SNP00124735	58	55	G	G	A	W14	n/a	n/a	n/a	n/a
58	7504768	4642229H1	SNP00037414	101	543	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4642923H1	SNP00046892	29	21	C	T	C	L3	0.06	0.04	0.04	0.03
58	7504768	4642923H1	SNP00124735	62	54	G	G	A	G14	n/a	n/a	n/a	n/a
58	7504768	4644309H1	SNP00043769	22	605	G	G	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4645524H1	SNP00046952	97	93	G	A	G	G27	n/a	n/a	n/a	n/a
58	7504768	4645524H1	SNP00054416	74	69	T	T	C	S19	n/a	n/a	n/a	n/a
58	7504768	4646632H1	SNP00146294	3	257	C	C	G	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	4646933H1	SNP00063753	19	13	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	4648313H1	SNP00063754	209	186	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	4648313H1	SNP00076561	141	118	G	G	C	G35	n/a	n/a	n/a	n/a
58	7504768	4648313H1	SNP00140832	236	213	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	4648743H1	SNP00037413	28	425	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	4818867H1	SNP00056531	124	639	C	C	A	noncoding	n/d	n/a	n/a	n/a
58	7504768	4818918H1	SNP00155250	233	368	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	4819335H1	SNP00063753	18	15	C	C	A	L1	n/a	n/a	n/a	n/a
58	7504768	4822515H1	SNP00037414	95	542	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4822515H1	SNP00054365	129	576	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	4822570H1	SNP00037414	81	553	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4822570H1	SNP00054365	115	587	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	4887930H1	SNP00129147	221	703	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	4888934H1	SNP00054416	45	42	T	T	C	L10	n/a	n/a	n/a	n/a
58	7504768	4889107H1	SNP00063753	25	18	C	C	A	R2	n/a	n/a	n/a	n/a
58	7504768	4890251H1	SNP00063833	128	132	G	G	C	A40	n/d	n/d	n/d	n/d
58	7504768	5003015H1	SNP00046952	109	101	G	A	G	S29	n/a	n/a	n/a	n/a
58	7504768	5003015H1	SNP00054416	86	78	T	T	C	F22	n/a	n/a	n/a	n/a
58	7504768	5003615H1	SNP00044426	98	246	G	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	5102495H1	SNP00155250	218	374	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	5102562H1	SNP00155250	218	375	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	5102562H1	SNP00044426	86	241	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	5102902H1	SNP00046952	114	107	G	A	G	Q31	n/a	n/a	n/a	n/a
58	7504768	5102902H1	SNP00054416	91	84	T	T	C	C24	n/a	n/a	n/a	n/a
58	7504768	5103232H1	SNP00037414	136	566	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	5103232H1	SNP00054365	170	600	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	5104276H1	SNP00146294	76	275	C	C	G	noncoding	n/a	n/a	n/a	n/a
58	7504768	5104476H1	SNP00052446	149	392	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	5106135H1	SNP00044426	41	240	A	G	A	noncoding	n/d	n/d	n/d	n/d

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	7504768	5107578H1	SNP00037414	49	546	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	5107578H1	SNP00054365	83	580	G	G	C	noncoding	n/a	n/a	n/a	n/a
58	7504768	5107620H1	SNP00044426	228	242	A	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	5108414H1	SNP00046952	76	74	G	A	G	A20	n/a	n/a	n/a	n/a
58	7504768	5764353H1	SNP00146294	343	372	G	C	G	noncoding	n/a	n/a	n/a	n/a
58	7504768	5765817H1	SNP00044426	298	298	G	G	A	noncoding	n/d	n/d	n/d	n/d
58	7504768	784132H1	SNP00011894	102	747	T	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	825709H1	SNP00052446	8	380	G	G	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	825709H1	SNP00067987	48	420	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	826297H1	SNP00054416	61	72	T	T	C	S20	n/a	n/a	n/a	n/a
58	7504768	829954H1	SNP00011894	160	697	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	829954H1	SNP00129147	141	678	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	832039H1	SNP00011894	189	705	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	832039H1	SNP00129147	170	686	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	842177H1	SNP00063753	8	4	C	C	A	noncoding	n/a	n/a	n/a	n/a
58	7504768	843350H1	SNP00037414	59	531	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	7504768	843350H1	SNP00054365	93	565	G	G	C	noncoding	n/a	n/a	n/a	n/a
59	7500757	1269986H1	SNP00028173	222	1745	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7500757	1475518H1	SNP00106848	126	1287	G	G	T	noncoding	n/d	n/d	n/d	n/d
59	7500757	2156540H1	SNP00106847	158	1145	C	C	G	noncoding	n/d	n/d	n/d	n/d
59	7500757	3915209H1	SNP00106846	111	872	C	T	C	noncoding	0.23	0.42	0.39	0.23
59	7500757	5773656H1	SNP00005518	250	1750	G	G	A	noncoding	n/a	n/a	n/a	n/a
59	7500757	8624396H1	SNP00154333	46	1864	T	T	C	noncoding	n/a	n/a	n/a	n/a
60	1730616	3188917H1	SNP00132428	171	255	C	T	C	A65	n/d	n/a	n/a	n/a
61	190404	3316961H1	SNP00104542	168	1543	A	A	G	K500	n/d	n/d	n/d	n/d
61	190404	5406235H1	SNP00057676	51	1324	G	T	G	G427	0.87	0.87	0.85	0.88
62	7500679	1235717H1	SNP00011316	78	1041	C	C	A	noncoding	0.99	0.69	0.88	0.82
62	7500679	1235717H1	SNP00011317	135	1098	G	C	G	noncoding	n/a	n/a	n/a	n/a
62	7500679	1285865H1	SNP00130638	183	389	A	A	G	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
62	7500679	1285865H1	SNP00132088	55	261	G	G	A	L68	n/a	n/a	n/a	n/a
63	7500687	1352080H1	SNP00011955	65	771	A	G	A	noncoding	0.12	n/a	n/a	n/a
63	7500687	1352080H1	SNP00044588	124	830	G	G	A	noncoding	n/d	n/a	n/a	n/a
63	7500687	1400156H1	SNP00132471	118	583	T	T	C	noncoding	n/a	n/a	n/a	n/a
63	7500687	2186481H1	SNP00044590	10	1187	T	T	C	noncoding	n/a	n/a	n/a	n/a
63	7500687	2269083H1	SNP00059865	249	965	G	A	G	noncoding	n/a	n/a	n/a	n/a
63	7500687	2593286H1	SNP00001723	127	221	A	A	G	P54	0.34	0.24	0.36	0.65
63	7500687	2697792H1	SNP00111346	74	349	A	G	A	Q97	n/a	n/a	n/a	n/a
63	7500687	4096988H1	SNP00044588	159	818	G	G	A	noncoding	n/d	n/a	n/a	n/a
63	7500687	4148849H1	SNP00044589	67	1119	G	A	G	noncoding	0.75	n/a	n/a	n/a
64	7500688	1352080H1	SNP00011955	65	1278	A	G	A	noncoding	0.12	n/a	n/a	n/a
64	7500688	1352080H1	SNP00044588	124	1337	G	G	A	noncoding	n/d	n/a	n/a	n/a
64	7500688	1400156H1	SNP00132471	118	1090	T	T	C	noncoding	n/a	n/a	n/a	n/a
64	7500688	2186481H1	SNP00044590	10	1694	T	T	C	noncoding	n/a	n/a	n/a	n/a
64	7500688	2269083H1	SNP00059865	249	1472	G	A	G	noncoding	n/a	n/a	n/a	n/a
64	7500688	2593286H1	SNP00001723	127	728	A	A	G	P223	0.34	0.24	0.36	0.65
64	7500688	2697792H1	SNP00111346	74	856	A	G	A	Q266	n/a	n/a	n/a	n/a
64	7500688	4096988H1	SNP00044588	159	1325	G	G	A	noncoding	n/d	n/a	n/a	n/a
64	7500688	4148849H1	SNP00044589	67	1626	G	A	G	noncoding	0.75	n/a	n/a	n/a
65	7500697	1339753H1	SNP00055204	163	405	C	C	A	noncoding	n/d	n/a	n/a	n/a
65	7500697	3535269H1	SNP00060602	96	668	C	C	T	noncoding	n/d	n/a	n/a	n/a
66	7500709	1402206H1	SNP00016820	9	2083	C	C	A	noncoding	n/d	n/a	n/a	n/a
66	7500709	1595983H1	SNP00016818	7	1943	C	C	T	R495	0.88	0.84	0.79	0.77
66	7500709	1989003H1	SNP00016822	166	2995	C	C	T	noncoding	n/a	n/a	n/a	n/a
66	7500709	2304976H1	SNP00098520	136	1637	A	A	C	T393	n/a	n/a	n/a	n/a
66	7500709	3074049H1	SNP00130277	79	384	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	7500709	3289688H1	SNP00140107	113	2611	C	C	T	noncoding	n/a	n/a	n/a	n/a
66	7500709	5771464H1	SNP00068738	455	1594	C	C	T	Y378	0.99	0.99	0.98	n/d
66	7500709	6338096H1	SNP00016821	431	2781	G	G	A	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
66	7500709	6481368H1	SNP00146224	398	713	G	G	C	D85	n/a	n/a	n/a	n/a
66	7500709	7059872H1	SNP00068738	175	1752	C	C	T	A431	0.99	0.99	0.98	n/d
66	7500709	7469363H1	SNP00152059	318	791	C	C	T	L111	n/a	n/a	n/a	n/a
67	7500711	1402206H1	SNP00016820	9	2212	C	C	A	noncoding	n/d	n/a	n/a	n/a
67	7500711	1595983H1	SNP00016818	7	2072	C	C	T	R538	0.88	0.84	0.79	0.77
67	7500711	1989003H1	SNP00016822	166	3124	C	C	T	noncoding	n/a	n/a	n/a	n/a
67	7500711	2304976H1	SNP00098520	136	1610	A	A	C	T384	n/a	n/a	n/a	n/a
67	7500711	2699021H1	SNP00068738	267	1723	C	C	T	H421	0.99	0.99	0.98	n/d
67	7500711	3074049H1	SNP00130277	79	384	A	A	G	noncoding	n/a	n/a	n/a	n/a
67	7500711	3289688H1	SNP00140107	113	2740	C	C	T	noncoding	n/a	n/a	n/a	n/a
67	7500711	6338096H1	SNP00016821	431	2910	G	G	A	noncoding	n/a	n/a	n/a	n/a
67	7500711	6481368H1	SNP00146224	398	713	G	G	C	D85	n/a	n/a	n/a	n/a
67	7500711	7469363H1	SNP00152059	318	791	C	C	T	L111	n/a	n/a	n/a	n/a
68	7500723	1005080H1	SNP00006328	84	147	T	C	T	L7	n/d	n/a	n/a	n/a
68	7500723	1220531H1	SNP00006329	203	1401	C	C	A	noncoding	n/a	n/a	n/a	n/a
68	7500723	1283008H1	SNP00055365	168	830	C	C	T	noncoding	0.99	n/a	n/a	n/a
68	7500723	1375166H1	SNP00006327	53	104	T	T	C	noncoding	n/d	n/a	n/a	n/a
68	7500723	1933092H1	SNP00029640	98	419	G	G	A	noncoding	n/a	n/a	n/a	n/a
69	7500764	1473474H1	SNP00010848	6	1489	C	C	T	noncoding	n/a	n/a	n/a	n/a
69	7500764	1553396H1	SNP00010849	174	1516	G	G	T	noncoding	n/a	n/a	n/a	n/a
69	7500764	1553396H1	SNP00148038	142	1484	C	C	T	noncoding	n/a	n/a	n/a	n/a
69	7500764	1995636H1	SNP00010847	71	1343	C	C	T	noncoding	n/a	n/a	n/a	n/a
69	7500764	2618624H1	SNP00149612	240	1209	C	C	A	noncoding	n/a	n/a	n/a	n/a
69	7500764	3435845H1	SNP00115242	111	1133	C	C	G	noncoding	n/d	n/d	n/d	n/d
69	7500764	6881314H1	SNP00124832	310	363	A	A	G	noncoding	n/a	n/a	n/a	n/a
71	7501350	1451293H1	SNP00007856	222	584	T	T	C	noncoding	n/a	n/a	n/a	n/a
71	7501350	1451293H1	SNP00014408	179	541	T	T	G	noncoding	n/a	n/a	n/a	n/a
72	7506396	1393846H1	SNP00036749	120	562	C	C	T	S135	n/a	n/a	n/a	n/a
72	7506396	1519711H1	SNP00074946	135	659	G	G	C	V167	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
72	7506396	1649353H1	SNP00074947	112	753	C	C	G	P199	n/a	n/a	n/a	n/a
72	7506396	2614262H1	SNP00026722	143	307	T	C	T	M50	n/d	n/d	n/d	n/d
73	7505917	1224269H1	SNP00019538	12	1406	C	C	G	G375	n/a	n/a	n/a	n/a
73	7505917	1224269H1	SNP00097325	89	1483	C	C	T	A401	n/a	n/a	n/a	n/a
73	7505917	1364234H1	SNP00097325	79	1484	C	C	T	G401	n/a	n/a	n/a	n/a
73	7505917	1467725H1	SNP00001985	87	1381	C	C	T	A367	n/a	n/a	n/a	n/a
73	7505917	1734934H1	SNP00006863	64	1175	C	C	T	P298	n/a	n/a	n/a	n/a
73	7505917	1734934H1	SNP00031912	83	1194	T	T	C	L305	n/a	n/a	n/a	n/a
73	7505917	2875519H1	SNP00031912	246	1193	C	T	C	A304	n/a	n/a	n/a	n/a
73	7505917	3408257H1	SNP00006863	195	1174	C	C	T	P298	n/a	n/a	n/a	n/a
73	7505917	3803327H1	SNP00006863	44	1172	C	C	T	L297	n/a	n/a	n/a	n/a
73	7505917	3803327H1	SNP00031912	63	1191	T	T	C	S304	n/a	n/a	n/a	n/a
73	7505917	3876468H1	SNP00019538	6	1404	C	C	G	R375	n/a	n/a	n/a	n/a
73	7505917	3876468H1	SNP00097325	84	1482	C	C	T	R401	n/a	n/a	n/a	n/a
73	7505917	3995861H1	SNP00019538	231	1396	C	C	G	A372	n/a	n/a	n/a	n/a
73	7505917	3997411H1	SNP00001985	214	1380	C	C	T	P367	n/a	n/a	n/a	n/a
73	7505917	3997411H1	SNP00019538	241	1405	C	C	G	A375	n/a	n/a	n/a	n/a
73	7505917	5992953H1	SNP00006863	138	1170	C	C	T	L297	n/a	n/a	n/a	n/a
73	7505917	5992953H1	SNP00031912	157	1189	T	T	C	V303	n/a	n/a	n/a	n/a
73	7505917	6307209H1	SNP00001985	326	1374	C	C	T	R365	n/a	n/a	n/a	n/a
73	7505917	6307209H1	SNP00031912	57	1109	T	T	C	A276	n/a	n/a	n/a	n/a
73	7505917	6307209H1	SNP00097325	433	1477	C	C	T	P399	n/a	n/a	n/a	n/a
73	7505917	6307313H1	SNP00006863	123	1095	C	C	T	L272	n/a	n/a	n/a	n/a
73	7505917	6307313H1	SNP00031912	142	1116	T	T	C	F279	n/a	n/a	n/a	n/a
73	7505917	6719860H1	SNP00001985	97	1461	T	C	T	stop394	n/a	n/a	n/a	n/a
73	7505917	6719860H1	SNP00019538	124	1494	C	C	G	P405	n/a	n/a	n/a	n/a
73	7505917	6719860H1	SNP00097325	202	1567	C	C	T	P429	n/a	n/a	n/a	n/a
73	7505917	6824111J1	SNP00063738	399	581	C	C	T	D100	0.67	0.50	0.74	0.64
73	7505917	7333804H1	SNP00097325	421	1575	T	C	T	S432	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
73	7505917	7398053H1	SNP00001985	432	1465	C	C	T	A395	n/a	n/a	n/a	n/a
74	7500701	1422265H1	SNP00020669	191	593	C	C	T	noncoding	n/a	n/a	n/a	n/a
74	7500701	2483938H1	SNP00022823	135	336	G	G	A	V94	n/d	n/a	n/a	n/a
74	7500701	3133415H1	SNP00022822	191	293	G	G	A	R80	n/d	n/d	0.96	n/d
75	7500702	1422265H1	SNP00020669	191	746	C	C	T	noncoding	n/a	n/a	n/a	n/a
75	7500702	2483938H1	SNP00022823	135	336	G	G	A	V94	n/d	n/a	n/a	n/a
75	7500702	3133415H1	SNP00022822	191	293	G	G	A	R80	n/d	n/d	0.96	n/d
76	6044343	1294191H1	SNP00022884	188	2771	T	T	G	noncoding	n/d	n/a	n/a	n/a
76	6044343	3449686H1	SNP00104924	146	1937	T	T	C	noncoding	n/d	n/a	n/a	n/a
76	6044343	4342984H1	SNP00022884	16	2768	T	T	G	noncoding	n/d	n/a	n/a	n/a
76	6044343	4839169H1	SNP00022884	4	2757	G	T	G	noncoding	n/d	n/a	n/a	n/a
76	6044343	6855931H1	SNP00022885	399	3235	G	G	A	noncoding	n/a	n/a	n/a	n/a
77	7503990	1341905H1	SNP00004157	64	3285	C	T	C	P913	0.46	n/a	n/a	n/a
77	7503990	1915202H1	SNP00115521	171	3947	T	T	C	noncoding	n/d	n/a	n/a	n/a
77	7503990	2021533H1	SNP00004158	87	4317	A	G	A	noncoding	0.73	0.91	n/d	0.78
77	7503990	2280153H1	SNP00024758	15	4215	T	T	C	noncoding	n/a	n/a	n/a	n/a
77	7503990	2600125H1	SNP00024755	83	2834	C	T	C	L763	0.72	0.93	n/d	0.77
77	7503990	2600125H1	SNP00024756	245	2996	T	C	T	Y817	0.78	n/a	n/a	n/a
77	7503990	2600125H1	SNP00024757	254	3005	T	C	T	F820	0.85	0.97	n/d	0.88
77	7503990	2736801H1	SNP00115519	14	1315	T	T	C	S256	n/d	n/d	n/d	n/d
77	7503990	3121639H1	SNP00115522	159	4062	C	C	T	noncoding	n/a	n/a	n/a	n/a
77	7503990	3228112H1	SNP00004157	10	3284	T	T	C	L913	0.46	n/a	n/a	n/a
77	7503990	3228112H1	SNP00096747	235	3509	T	T	C	F988	0.77	0.92	0.99	0.82
77	7503990	3270946H1	SNP00004157	50	3282	T	T	C	I912	0.46	n/a	n/a	n/a
77	7503990	3661149H1	SNP00115521	172	3945	T	T	C	noncoding	n/d	n/a	n/a	n/a
77	7503990	3661149H1	SNP00115522	286	4060	C	C	T	noncoding	n/a	n/a	n/a	n/a
77	7503990	3679593H1	SNP00115521	142	3942	T	T	C	noncoding	n/d	n/a	n/a	n/a
77	7503990	3679593H1	SNP00115522	257	4057	C	C	T	noncoding	n/a	n/a	n/a	n/a
77	7503990	3681589H1	SNP00115522	100	4064	C	C	T	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
77	7503990	3917071H1	SNP00096747	91	3507	T	T	C	L987	0.77	0.92	0.99	0.82
77	7503990	4163704H1	SNP00151675	128	1209	C	C	T	T221	n/a	n/a	n/a	n/a
77	7503990	4183129H1	SNP00004158	71	4316	A	G	A	noncoding	0.73	0.91	n/d	0.78
77	7503990	4348376H1	SNP00004158	134	4306	A	G	A	noncoding	0.73	0.91	n/d	0.78
77	7503990	4348376H1	SNP00024758	32	4204	C	T	C	noncoding	n/a	n/a	n/a	n/a
77	7503990	4455065H1	SNP00096747	58	3510	T	T	C	I988	0.77	0.92	0.99	0.82
77	7503990	4553662H1	SNP00024758	197	4213	T	T	C	noncoding	n/a	n/a	n/a	n/a
77	7503990	4571868H1	SNP00004157	113	3283	C	T	C	T912	0.46	n/a	n/a	n/a
77	7503990	5026526H1	SNP00004158	134	4304	G	G	A	noncoding	0.73	0.91	n/d	0.78
77	7503990	5026526H1	SNP00024758	32	4202	T	T	C	noncoding	n/a	n/a	n/a	n/a
77	7503990	5046938H1	SNP00062181	84	959	G	G	T	V138	n/a	n/a	n/a	n/a
77	7503990	5223889H1	SNP00024758	62	4216	T	T	C	noncoding	n/a	n/a	n/a	n/a
77	7503990	5726796H1	SNP00004158	107	4252	G	G	A	noncoding	0.73	0.91	n/d	0.78
77	7503990	645839H1	SNP00024758	40	4193	T	T	C	noncoding	n/a	n/a	n/a	n/a
77	7503990	6549708H1	SNP00115520	11	1939	T	T	C	R464	n/d	n/d	n/d	n/d
77	7503990	6573209H1	SNP00092898	173	2597	T	T	C	S684	0.79	0.95	n/d	0.86
77	7503990	6804476J1	SNP00092496	367	1558	C	C	G	S337	0.87	n/a	0.61	0.60
77	7503990	7152502H1	SNP00092496	317	1445	C	C	G	R300	0.87	n/a	0.61	0.60
77	7503990	7161660H1	SNP00062181	208	961	G	G	T	L138	n/a	n/a	n/a	n/a
77	7503990	7262641H1	SNP00024755	269	2836	T	T	C	F763	0.72	0.93	n/d	0.77
77	7503990	7262641H1	SNP00024756	431	2998	C	C	T	D817	0.78	n/a	n/a	n/a
77	7503990	7262641H1	SNP00024757	440	3007	C	C	T	F820	0.85	0.97	n/d	0.88
77	7503990	7326295H1	SNP00024758	84	4196	T	T	C	noncoding	n/a	n/a	n/a	n/a
77	7503990	7326295H1	SNP00115522	237	4043	C	C	T	noncoding	n/a	n/a	n/a	n/a
77	7503990	7603261H1	SNP00115521	504	3956	T	T	C	noncoding	n/d	n/a	n/a	n/a
77	7503990	8626011J1	SNP00062181	706	981	T	G	T	V145	n/a	n/a	n/a	n/a
77	7503990	968309H1	SNP00024758	110	4211	T	T	C	noncoding	n/a	n/a	n/a	n/a
78	7504655	6395382H1	SNP00000037	271	910	T	T	C	noncoding	n/a	n/a	n/a	n/a
79	7504690	1261613H1	SNP00044450	129	872	C	C	T	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
79	7504690	1405907H1	SNP00044451	83	1414	T	C	T	noncoding	n/a	n/a	n/a	n/a
79	7504690	1680256H1	SNP00044450	176	869	C	C	T	noncoding	n/a	n/a	n/a	n/a
79	7504690	2040386H1	SNP00044451	65	1413	C	C	T	noncoding	n/a	n/a	n/a	n/a
79	7504690	5575755H1	SNP00044450	119	871	C	C	T	noncoding	n/a	n/a	n/a	n/a
79	7504690	6397357H1	SNP00044450	24	865	C	C	T	noncoding	n/a	n/a	n/a	n/a
80	7504720	025832H1	SNP00006719	154	992	A	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	1000318H1	SNP00070116	15	33	T	T	C	noncoding	n/a	n/a	n/a	n/a
80	7504720	1349406H1	SNP00144339	210	904	A	A	G	noncoding	n/a	n/a	n/a	n/a
80	7504720	138512H1	SNP00070116	27	32	T	T	C	noncoding	n/a	n/a	n/a	n/a
80	7504720	1464560H1	SNP00006719	69	986	G	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	167768H1	SNP00070116	28	21	T	T	C	noncoding	n/a	n/a	n/a	n/a
80	7504720	1932685H1	SNP00006719	242	993	G	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	1985942H1	SNP00144339	212	905	A	A	G	noncoding	n/a	n/a	n/a	n/a
80	7504720	1997242H1	SNP00006719	184	965	G	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	2013940H1	SNP00144339	194	897	A	A	G	noncoding	n/a	n/a	n/a	n/a
80	7504720	2108806H1	SNP00006719	166	982	G	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	2968021H1	SNP00070116	17	31	T	T	C	noncoding	n/a	n/a	n/a	n/a
80	7504720	3236115H1	SNP00006719	220	990	G	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	3284457H1	SNP00070116	17	30	C	T	C	noncoding	n/a	n/a	n/a	n/a
80	7504720	3370771H1	SNP00006719	151	991	A	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	3472889H1	SNP00070116	23	27	C	T	C	noncoding	n/a	n/a	n/a	n/a
80	7504720	3607738H1	SNP00006719	125	987	G	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	3705874H1	SNP00006719	270	989	A	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	3735181H1	SNP00144339	212	896	G	A	G	noncoding	n/a	n/a	n/a	n/a
80	7504720	3811868H1	SNP00006719	193	1000	G	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	3860814H1	SNP00144339	231	894	A	A	G	noncoding	n/a	n/a	n/a	n/a
80	7504720	4334301H1	SNP00006719	227	981	G	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	6722906H1	SNP00006719	334	983	A	G	A	noncoding	n/a	n/a	n/a	n/a
80	7504720	688061H1	SNP00006719	137	985	G	G	A	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
81	7504722	030692H1	SNP00124454	96	813	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	1286313H1	SNP00092908	92	1122	G	G	A	noncoding	n/a	n/a	n/a	n/a
81	7504722	154958H1	SNP00018343	135	558	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	2399811H1	SNP00124454	11	811	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	3109003H1	SNP00124454	138	812	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	3158872H1	SNP00124454	162	809	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	4108913H1	SNP00124454	176	805	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	4211203H1	SNP00124454	182	810	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	4531755H1	SNP00092908	84	1119	G	G	A	noncoding	n/a	n/a	n/a	n/a
81	7504722	5699517H1	SNP00124454	195	807	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	6737674H1	SNP00124454	149	764	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	6929007H1	SNP00018343	366	557	C	C	T	noncoding	n/a	n/a	n/a	n/a
81	7504722	879871H1	SNP00092908	49	1120	G	G	A	noncoding	n/a	n/a	n/a	n/a
82	7504733	025850H1	SNP00140932	140	289	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	026018H1	SNP00034039	94	239	C	C	T	L19	n/a	n/a	n/a	n/a
82	7504733	026018H1	SNP00140932	141	290	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	138450H1	SNP00034039	90	223	C	C	T	I13	n/a	n/a	n/a	n/a
82	7504733	138450H1	SNP00140932	137	270	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	139638H1	SNP00034039	92	238	C	C	T	V18	n/a	n/a	n/a	n/a
82	7504733	139638H1	SNP00140932	139	288	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	139751H1	SNP00034039	71	209	C	C	T	L9	n/a	n/a	n/a	n/a
82	7504733	166203H1	SNP00034039	61	234	C	C	T	A17	n/a	n/a	n/a	n/a
82	7504733	166203H1	SNP00140932	108	284	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	166663H1	SNP00034039	88	237	C	C	T	A18	n/a	n/a	n/a	n/a
82	7504733	166663H1	SNP00140932	135	287	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	166896H1	SNP00034039	88	228	C	C	T	T15	n/a	n/a	n/a	n/a
82	7504733	166896H1	SNP00140932	135	277	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	167123H1	SNP00034039	87	236	C	C	T	L18	n/a	n/a	n/a	n/a
82	7504733	167123H1	SNP00140932	134	286	T	T	C	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
82	7504733	2048217H1	SNP00034041	53	387	T	T	G	noncoding	n/a	n/a	n/a	n/a
82	7504733	2513378H1	SNP00034039	76	235	C	C	T	D17	n/a	n/a	n/a	n/a
82	7504733	2513378H1	SNP00140932	123	285	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	2516759H1	SNP00034039	92	240	C	C	T	P19	n/a	n/a	n/a	n/a
82	7504733	2516759H1	SNP00140932	139	291	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	271049H1	SNP00034039	61	225	C	C	T	S14	n/a	n/a	n/a	n/a
82	7504733	271049H1	SNP00140932	108	272	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	272900H1	SNP00034039	76	257	C	C	T	noncoding	n/a	n/a	n/a	n/a
82	7504733	272900H1	SNP00140932	123	306	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	293166H1	SNP00034039	74	227	C	C	T	R15	n/a	n/a	n/a	n/a
82	7504733	293166H1	SNP00140932	121	275	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	293308H1	SNP00034039	70	217	C	C	T	L11	n/a	n/a	n/a	n/a
82	7504733	293308H1	SNP00140932	117	267	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	294628H1	SNP00034039	69	222	C	C	T	T13	n/a	n/a	n/a	n/a
82	7504733	294952H1	SNP00034041	165	366	T	T	G	noncoding	n/a	n/a	n/a	n/a
82	7504733	2957315H1	SNP00034041	208	386	T	T	G	noncoding	n/a	n/a	n/a	n/a
82	7504733	4086311H1	SNP00034041	217	384	T	T	G	noncoding	n/a	n/a	n/a	n/a
82	7504733	4416404H1	SNP00034041	225	385	T	T	G	noncoding	n/a	n/a	n/a	n/a
82	7504733	4797147H1	SNP00034039	90	231	C	C	T	P16	n/a	n/a	n/a	n/a
82	7504733	4797147H1	SNP00140932	137	279	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	480396H1	SNP00034040	152	369	A	A	G	noncoding	n/a	n/a	n/a	n/a
82	7504733	5951186H1	SNP00140932	136	283	T	T	C	noncoding	n/a	n/a	n/a	n/a
82	7504733	7024134H1	SNP00140932	156	297	T	T	C	noncoding	n/a	n/a	n/a	n/a
83	7507100	1560345H1	SNP00047855	58	1891	T	T	C	noncoding	0.59	0.61	0.54	0.69
84	7503330	3783254H1	SNP00124674	59	70	A	A	G	K8	n/a	n/a	n/a	n/a
85	7504519	2432516H1	SNP00026362	61	74	A	A	G	noncoding	n/a	n/a	n/a	n/a
85	7504519	2432609H1	SNP00056302	150	1907	A	A	G	noncoding	n/d	n/a	n/a	n/a
85	7504519	2655227H1	SNP00131351	257	2843	C	T	C	noncoding	n/a	n/a	n/a	n/a
85	7504519	3223832H1	SNP00131351	221	2841	C	T	C	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
85	7504519	3223832H1	SNP00131352	271	2891	C	C	A	noncoding	n/a	n/a	n/a	n/a
85	7504519	7169295H1	SNP00131352	437	2893	C	C	A	noncoding	n/a	n/a	n/a	n/a
87	7504738	010054H1	SNP00061272	134	393	A	A	C	noncoding	n/a	n/a	n/a	n/a
87	7504738	010089H1	SNP00058947	176	73	C	C	T	L3	n/a	n/a	n/a	n/a
87	7504738	010482H1	SNP00146399	254	159	G	G	A	E31	n/a	n/a	n/a	n/a
87	7504738	010838H1	SNP00002068	14	31	A	A	C	noncoding	n/a	n/a	n/a	n/a
87	7504738	010934H1	SNP00019749	38	55	T	G	T	noncoding	n/a	n/a	n/a	n/a
87	7504738	010934H1	SNP00149440	83	100	A	G	A	M12	n/a	n/a	n/a	n/a
87	7504738	1289761H1	SNP00149440	73	99	G	G	A	L11	n/a	n/a	n/a	n/a
87	7504738	154781H1	SNP00002068	14	30	A	A	C	noncoding	n/a	n/a	n/a	n/a
87	7504738	155182H1	SNP00061272	131	379	A	A	C	noncoding	n/a	n/a	n/a	n/a
87	7504738	2228722H1	SNP00149442	50	152	C	C	T	P29	n/a	n/a	n/a	n/a
87	7504738	2439274H1	SNP00061272	63	394	A	A	C	noncoding	n/a	n/a	n/a	n/a
87	7504738	3162461H1	SNP00058947	192	70	C	C	T	Q2	n/a	n/a	n/a	n/a
87	7504738	3229391H1	SNP00061272	30	398	A	A	C	noncoding	n/a	n/a	n/a	n/a
87	7504738	3294683H1	SNP00058947	191	72	C	C	T	N2	n/a	n/a	n/a	n/a
87	7504738	3685171H1	SNP00146399	272	160	G	G	A	A32	n/a	n/a	n/a	n/a
87	7504738	3844355H1	SNP00052991	123	235	C	C	T	noncoding	n/d	n/a	n/a	n/a
87	7504738	3845160H1	SNP00146399	274	158	G	G	A	G31	n/a	n/a	n/a	n/a
87	7504738	4345360H1	SNP00052991	86	230	C	C	T	noncoding	n/d	n/a	n/a	n/a
87	7504738	4345360H1	SNP00146399	13	157	G	G	A	E31	n/a	n/a	n/a	n/a
87	7504738	4345451H1	SNP00058947	163	71	C	C	T	T2	n/a	n/a	n/a	n/a
87	7504738	4347872H1	SNP00058947	151	69	C	C	T	I1	n/a	n/a	n/a	n/a
87	7504738	4994751H1	SNP00002068	27	28	A	A	C	noncoding	n/a	n/a	n/a	n/a
87	7504738	560674H1	SNP00052991	118	231	C	C	T	noncoding	n/d	n/a	n/a	n/a
87	7504738	5978392H1	SNP00058947	178	63	C	C	T	noncoding	n/a	n/a	n/a	n/a
87	7504738	5978861H1	SNP00058947	172	74	C	C	T	P3	n/a	n/a	n/a	n/a
87	7504738	5982159H1	SNP00058947	185	67	C	C	T	L1	n/a	n/a	n/a	n/a
87	7504738	5986062H1	SNP00146399	259	154	G	G	A	E30	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
87	7504738	6091956H1	SNP00002069	169	60	T	T	C	noncoding	n/d	0.01	0.06	n/d
88	7510280	3735009H1	SNP00105730	81	381	C	C	T	noncoding	n/a	n/a	n/a	n/a
88	7510280	8103858H1	SNP00105730	143	393	C	C	T	noncoding	n/a	n/a	n/a	n/a
89	7503700	1336111H1	SNP00151238	61	1137	A	A	G	K239	n/a	n/a	n/a	n/a
89	7503700	3417040H1	SNP00151239	194	1388	C	C	T	noncoding	n/a	n/a	n/a	n/a
89	7503700	4596758H1	SNP00151238	143	1138	A	A	G	K239	n/a	n/a	n/a	n/a
89	7503700	5943105H1	SNP00151239	45	1386	C	C	T	noncoding	n/a	n/a	n/a	n/a
90	7504685	2483756H1	SNP00000806	335	586	G	G	A	noncoding	n/a	n/a	n/a	n/a
90	7504685	6119436H1	SNP00000806	192	548	A	G	A	noncoding	n/a	n/a	n/a	n/a
90	7504685	6119636H1	SNP00000806	220	587	A	G	A	noncoding	n/a	n/a	n/a	n/a
91	7506844	1456284H1	SNP00140641	213	437	C	C	A	noncoding	n/a	n/a	n/a	n/a
91	7506844	1541920H1	SNP00140641	48	430	C	C	A	noncoding	n/a	n/a	n/a	n/a
91	7506844	1622811H1	SNP00140641	35	438	C	C	A	noncoding	n/a	n/a	n/a	n/a
91	7506844	2110750H1	SNP00053655	125	724	G	G	A	noncoding	n/a	n/a	n/a	n/a
91	7506844	2138747H1	SNP00050861	186	136	G	G	T	noncoding	0.99	n/a	n/a	n/a
91	7506844	2322502H1	SNP00140641	54	434	C	C	A	noncoding	n/a	n/a	n/a	n/a
91	7506844	4527351H1	SNP00050861	178	131	G	G	T	noncoding	0.99	n/a	n/a	n/a
91	7506844	4639711H1	SNP00050861	86	134	G	G	T	noncoding	0.99	n/a	n/a	n/a
91	7506844	4946835H1	SNP00140641	126	369	C	C	A	noncoding	n/a	n/a	n/a	n/a
91	7506844	5163294H1	SNP00050861	52	135	G	G	T	noncoding	0.99	n/a	n/a	n/a
91	7506844	5165765H1	SNP00050861	173	127	G	G	T	noncoding	0.99	n/a	n/a	n/a
91	7506844	5533895H1	SNP00140641	193	436	C	C	A	noncoding	n/a	n/a	n/a	n/a
96	7503772	1212254H1	SNP00036241	31	4506	C	C	T	noncoding	0.05	n/a	n/a	n/a
96	7503772	1212254H1	SNP00036242	172	4647	G	G	A	noncoding	n/a	n/a	n/a	n/a
96	7503772	1215514H1	SNP00036243	200	4772	G	G	C	noncoding	n/a	n/a	n/a	n/a
96	7503772	2611731H1	SNP00036239	75	4252	C	C	A	noncoding	n/a	n/a	n/a	n/a
96	7503772	2611731H1	SNP00036240	111	4288	C	C	T	noncoding	n/a	n/a	n/a	n/a
96	7503772	3087123H1	SNP00036242	77	4645	G	G	A	noncoding	n/a	n/a	n/a	n/a
96	7503772	3087123H1	SNP00036243	202	4770	G	G	C	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
96	7503772	3286120H2	SNP00036239	214	4254	C	C	A	noncoding	n/a	n/a	n/a	n/a
96	7503772	3286120H2	SNP00036240	250	4291	C	C	T	noncoding	n/a	n/a	n/a	n/a
96	7503772	3372558H1	SNP00036243	81	4763	G	G	C	noncoding	n/a	n/a	n/a	n/a
96	7503772	3373510H1	SNP00036243	173	4767	G	G	C	noncoding	n/a	n/a	n/a	n/a
96	7503772	3565279H1	SNP00036240	167	4290	T	C	T	noncoding	n/a	n/a	n/a	n/a
96	7503772	4188096H1	SNP00036242	36	4643	G	G	A	noncoding	n/a	n/a	n/a	n/a
96	7503772	4466544H1	SNP00036239	177	4255	C	C	A	noncoding	n/a	n/a	n/a	n/a
96	7503772	4466544H1	SNP00036240	213	4292	C	C	T	noncoding	n/a	n/a	n/a	n/a
96	7503772	4895923H1	SNP00036243	235	4769	G	G	C	noncoding	n/a	n/a	n/a	n/a
96	7503772	6493623H1	SNP00036241	84	3570	T	C	T	noncoding	0.05	n/a	n/a	n/a
96	7503772	6531015H1	SNP00036242	45	4633	G	G	A	noncoding	n/a	n/a	n/a	n/a
96	7503772	6531015H1	SNP00036243	170	4758	G	G	C	noncoding	n/a	n/a	n/a	n/a
96	7503772	6544492H1	SNP00036241	182	4504	C	C	T	noncoding	0.05	n/a	n/a	n/a
96	7503772	7037312H1	SNP00036239	329	2389	C	C	A	P762	n/a	n/a	n/a	n/a
96	7503772	7037312H1	SNP00036240	365	2427	C	C	T	H774	n/a	n/a	n/a	n/a
96	7503772	7626412J1	SNP00138319	187	3414	T	T	C	noncoding	n/a	n/a	n/a	n/a
96	7503772	8037780H1	SNP00138319	139	1718	C	T	C	P538	n/a	n/a	n/a	n/a
97	7503773	1212254H1	SNP00036241	31	4459	C	C	T	noncoding	0.05	n/a	n/a	n/a
97	7503773	1212254H1	SNP00036242	172	4600	G	G	A	noncoding	n/a	n/a	n/a	n/a
97	7503773	1215514H1	SNP00036243	200	4725	G	G	C	noncoding	n/a	n/a	n/a	n/a
97	7503773	2611731H1	SNP00036239	75	4205	C	C	A	noncoding	n/a	n/a	n/a	n/a
97	7503773	2611731H1	SNP00036240	111	4241	C	C	T	noncoding	n/a	n/a	n/a	n/a
97	7503773	3087123H1	SNP00036242	77	4598	G	G	A	noncoding	n/a	n/a	n/a	n/a
97	7503773	3087123H1	SNP00036243	202	4723	G	G	C	noncoding	n/a	n/a	n/a	n/a
97	7503773	3286120H2	SNP00036239	214	4207	C	C	A	noncoding	n/a	n/a	n/a	n/a
97	7503773	3286120H2	SNP00036240	250	4244	C	C	T	noncoding	n/a	n/a	n/a	n/a
97	7503773	3372558H1	SNP00036243	81	4716	G	G	C	noncoding	n/a	n/a	n/a	n/a
97	7503773	3373510H1	SNP00036243	173	4720	G	G	C	noncoding	n/a	n/a	n/a	n/a
97	7503773	3565279H1	SNP00036240	167	4243	T	C	T	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
97	7503773	4188096H1	SNP00036242	36	4596	G	G	A	noncoding	n/a	n/a	n/a	n/a
97	7503773	4466544H1	SNP00036239	177	4208	C	C	A	noncoding	n/a	n/a	n/a	n/a
97	7503773	4466544H1	SNP00036240	213	4245	C	C	T	noncoding	n/a	n/a	n/a	n/a
97	7503773	4895923H1	SNP00036243	235	4722	G	G	C	noncoding	n/a	n/a	n/a	n/a
97	7503773	6493623H1	SNP00036241	84	3652	T	C	T	C1183	0.05	n/a	n/a	n/a
97	7503773	6531015H1	SNP00036242	45	4586	G	G	A	noncoding	n/a	n/a	n/a	n/a
97	7503773	6531015H1	SNP00036243	170	4711	G	G	C	noncoding	n/a	n/a	n/a	n/a
97	7503773	6544492H1	SNP00036241	182	4457	C	C	T	noncoding	0.05	n/a	n/a	n/a
97	7503773	8037780H1	SNP00138319	139	1718	C	T	C	P538	n/a	n/a	n/a	n/a
98	7504698	030732H1	SNP00149756	215	453	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	031619H1	SNP00149756	216	451	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	072413H1	SNP00149756	118	449	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	2188139H1	SNP00149756	248	454	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	2723635H1	SNP00003052	35	527	G	G	A	noncoding	1.00	n/a	n/a	n/a
98	7504698	3090344H1	SNP00149756	102	452	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	3572162H1	SNP00003052	50	526	G	G	A	noncoding	1.00	n/a	n/a	n/a
98	7504698	3806693H1	SNP00149756	93	448	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	4071296H1	SNP00003052	27	525	G	G	A	noncoding	1.00	n/a	n/a	n/a
98	7504698	5276116H1	SNP00003052	24	520	G	G	A	noncoding	1.00	n/a	n/a	n/a
98	7504698	5768664H1	SNP00141546	431	936	C	C	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	5949629H1	SNP00149756	250	450	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	6211253H1	SNP00149756	73	426	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7504698	6938844H1	SNP00003052	289	421	G	G	A	noncoding	1.00	n/a	n/a	n/a
99	7510361	1325602H1	SNP00105072	246	666	T	T	C	W207	n/a	n/a	n/a	n/a
99	7510361	1326305H1	SNP00137858	104	460	C	C	T	P138	n/a	n/a	n/a	n/a
99	7510361	1824780H1	SNP00105070	252	252	A	A	G	S69	n/d	n/d	n/d	n/d
99	7510361	1824858H1	SNP00105071	167	525	A	A	G	T160	n/d	n/d	n/d	n/d
99	7510361	1824858H1	SNP00105072	242	600	T	T	C	L185	n/a	n/a	n/a	n/a
99	7510361	1825455H1	SNP00105072	195	665	T	T	C	L206	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
99	7510361	7653234H1	SNP00111130	13	707	A	A	G	E220	n/d	n/d	n/d	n/d
99	7510361	7939670H1	SNP00105071	295	589	A	A	G	N181	n/d	n/d	n/d	n/d
99	7510361	7939670H1	SNP00105072	370	667	C	T	C	S207	n/a	n/a	n/a	n/a
100	7507013	6041339H1	SNP00058963	323	723	C	T	C	noncoding	n/d	n/d	n/a	n/d
101	7510507	1748724H1	SNP00151312	100	2896	A	A	G	noncoding	n/a	n/a	n/a	n/a
101	7510507	2137158H1	SNP00031810	137	1955	C	C	T	noncoding	0.99	n/a	n/a	n/a
101	7510507	2864359H1	SNP00039104	136	878	C	C	T	G198	n/a	n/a	n/a	n/a
101	7510507	3327158H1	SNP00039104	242	879	C	C	T	P199	n/a	n/a	n/a	n/a
101	7510507	4376488H1	SNP00039104	204	876	T	C	T	C198	n/a	n/a	n/a	n/a
101	7510507	5691218H1	SNP00151312	247	2894	A	A	G	noncoding	n/a	n/a	n/a	n/a
101	7510507	6409380H1	SNP00151312	177	2855	A	A	G	noncoding	n/a	n/a	n/a	n/a
101	7510507	6482421H1	SNP00039104	288	880	C	C	T	T199	n/a	n/a	n/a	n/a
101	7510507	6623018J1	SNP00138949	98	1187	T	T	C	F301	n/a	n/a	n/a	n/a

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51,

b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, SEQ ID NO:7, SEQ ID NO:10-11, SEQ ID NO:14, SEQ ID NO:17-20, SEQ ID NO:23-25, SEQ ID NO:27-38, SEQ ID NO:40-46, and SEQ ID NO:48-51

c) a polypeptide comprising a naturally occurring amino acid sequence at least 92% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:39, and SEQ ID NO:47,

d) a polypeptide comprising a naturally occurring amino acid sequence at least 93% identical to the amino acid sequence of SEQ ID NO:21,

e) a polypeptide comprising a naturally occurring amino acid sequence at least 94% identical to the amino acid sequence of SEQ ID NO:9,

f) a polypeptide comprising a naturally occurring amino acid sequence at least 98% identical to the amino acid sequence of SEQ ID NO:13,

g) a polypeptide consisting essentially of a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:12, SEQ ID NO:15, SEQ ID NO:22, SEQ ID NO:26,

h) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, and

i) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-51.

2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method of producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-51.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-102,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:52-53, SEQ ID NO:59-61, SEQ ID NO:65-67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73-75, SEQ ID NO:77-79, SEQ ID NO:85-86, SEQ ID NO:89-92, SEQ ID NO:95-97, and SEQ ID NO:99-102,

c) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 91% identical to a polynucleotide sequence selected from the group consisting of SEQ

ID NO:62, and SEQ ID NO:88,

d) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 93% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:54-56, SEQ ID NO:63, SEQ ID NO:83, and SEQ ID NO:93,

e) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 94% identical to the polynucleotide sequence of SEQ ID NO:64,

f) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 96% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:68, SEQ ID NO:81, and SEQ ID NO:87,

g) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 98% identical to the polynucleotide sequence of SEQ ID NO:94,

h) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 99% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:57, SEQ ID NO:82, and SEQ ID NO:98,

i) a polynucleotide consisting essentially of a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:58, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:76, SEQ ID NO:80, and SEQ ID NO:84,

j) a polynucleotide complementary to a polynucleotide of a),

k) a polynucleotide complementary to a polynucleotide of b),

l) a polynucleotide complementary to a polynucleotide of c),

m) a polynucleotide complementary to a polynucleotide of d),

n) a polynucleotide complementary to a polynucleotide of e),

o) a polynucleotide complementary to a polynucleotide of f),

p) a polynucleotide complementary to a polynucleotide of g),

q) a polynucleotide complementary to a polynucleotide of h),

r) a polynucleotide complementary to a polynucleotide of i), and

s) an RNA equivalent of a)-r).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-51.

19. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test

compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

a) treating a biological sample containing nucleic acids with the test compound,

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

c) quantifying the amount of hybridization complex, and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A method for a diagnostic test for a condition or disease associated with the expression of SECP in a biological sample, the method comprising:

a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex,

and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, further comprising a label.

35. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibodies from the animal, and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51.

37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-51, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibody producing cells from the animal,

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,

d) culturing the hybridoma cells, and

e) isolating from the culture monoclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51 in a sample, the method comprising:

a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51 from a sample, the method comprising:

a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-51.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

92. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

93. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

94. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

95. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

96. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

97. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

98. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.

99. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.

100. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:45.

101. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:46.

102. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:47.

103. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:48.

104. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:49.

105. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:50.

106. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:51.

107. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:52.

108. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:53.

109. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:54.

110. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:55.

111. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:56.

112. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:57.

113. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:58.

114. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:59.

115. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:60.

116. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:61.

117. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:62.

118. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:63.

119. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:64.

120. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:65.

121. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:66.

122. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:67.

123. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:68.

124. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:69.

125. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:70.

126. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:71.

127. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:72.

128. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:73.

129. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:74.

130. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:75.

131. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:76.

132. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:77.

133. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:78.

134. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:79.

135. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:80.

136. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:81.

137. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:82.

138. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:83.

139. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:84.

140. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:85.

141. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:86.

142. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:87.

143. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:88.

144. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:89.

145. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:90.

146. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:91.

147. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:92.

148. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:93.

149. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:94.

150. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:95.

151. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:96.

152. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:97.

153. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:98.

154. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:99.

155. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:100.

156. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:101.

157. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:102.

<110> INCYTE GENOMICS, INC.
LEHR-MASON, Patricia M.
KABLE, Amy E.
ELLIOTT, Vicki S.
MARQUIS, Joseph P.
BAUGHN, Mariah R.
CHAWLA, Narinder K.
TRAN, Uyen K.
JIN, Pei
TANG, Y. Tom
ZEBARJADIAN, Yeganeh
SWARNAKAR, Anita
HAFALIA, April J.A.
COCKS, Benjamin G.
WARREN, Bridget A.
EMERLING, Brooke M.
PEARSON, Cecilia I.
CHIEN, David
PETERSON, David P.
FU, Glenn K.
YUE, Henry
JACKSON, Alan A.
JIANG, Xin
HAWKINS, Phillip R.
LAL, Preeti G.
KHARE, Reena
LEE, Sally
LEE, Soo Yeun
RICHARDSON, Thomas W.
CHANG, Hsin-Ru

<120> SECRETED PROTEINS

<130> PF-1384 PCT

<140> To Be Assigned
<141> Herewith

<150> US 60/357,002
<151> 2002-02-13

<150> US 60/362,439
<151> 2002-03-06

<150> US 60/366,041
<151> 2002-03-19

<160> 102
<170> PERL Program

<210> 1
<211> 120
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 7506904CD1

<400> 1
Met Gly Gly Tyr Leu Lys Thr Arg Pro Trp Thr Leu Gln His Phe
1 5 10 15
Tyr Leu Cys Leu Met Pro Ala Ala Thr Trp Leu Val Leu Leu Leu
20 25 30
Leu Leu Trp Leu Ser Leu Gly Val Lys Thr Gly Ser Cys Ser Gln
35 40 45
Pro Gln Asn Leu Cys Cys Leu Gly Thr Asp His His Cys Lys Arg
50 55 60
Gly Ser Cys Tyr Cys Asp Glu Phe Cys His Val Ala Pro Asp Cys
65 70 75
His Pro Asp His Ser Val Leu Cys Asn Pro Gly Asn Ser His Thr
80 85 90
Gly Pro Ile Pro Pro Thr Ala Lys Leu Asp Ala Met Ala Gly Arg
95 100 105
Gly Lys Pro Phe Ala Cys Thr Ser Gly Gln Ser Arg Asp Val Ala
110 115 120

<210> 2
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506909CD1

<400> 2
Met Gly Gly Tyr Leu Lys Thr Arg Pro Trp Thr Leu Gln His Phe
1 5 10 15
Tyr Leu Cys Leu Met Pro Ala Ala Thr Trp Leu Val Leu Leu Leu
20 25 30
Leu Leu Trp Leu Ser Leu Gly Val Lys Thr Ala Ser Gln Met Thr
35 40 45
Lys Met Val Leu Gln Met Val Leu Arg Met Glu Asn Pro Pro Ser
50 55 60
Pro Ala Arg Ser His Leu Asp Trp Met Gln Ser Met Val Ser Ser
65 70 75
Leu Gln Val Leu

<210> 3
<211> 270
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507096CD1

<400> 3
Met Ala Thr Leu Ser Phe Val Phe Leu Leu Leu Gly Ala Val Ser

```
  1               5                  10                  15
Trp Pro Pro Ala Ser Ala Ser Gly Gln Glu Phe Trp Pro Gly Gln
                 20                  25                  30
Ser Ala Ala Asp Ile Leu Ser Gly Ala Ala Ser Arg Arg Arg Tyr
                 35                  40                  45
Leu Leu Tyr Asp Val Asn Pro Pro Glu Gly Phe Asn Leu Arg Arg
                 50                  55                  60
Asp Val Tyr Ile Arg Ile Ala Ser Leu Leu Lys Thr Leu Leu Lys
                 65                  70                  75
Thr Glu Glu Trp Val Leu Val Leu Pro Pro Trp Gly Arg Leu Tyr
                 80                  85                  90
His Trp Gln Ser Pro Asp Ile His Gln Val Arg Ile Pro Trp Ser
                 95                 100                 105
Glu Phe Phe Asp Leu Pro Ser Leu Asn Lys Asn Ile Pro Val Ile
                110                 115                 120
Glu Tyr Glu Gln Phe Ile Ala Gly Leu Ser Leu His Arg Gly Ala
                125                 130                 135
Pro Ala Ala Glu Lys His Ile Ser Pro Val Arg Asp Val Arg Gln
                140                 145                 150
Ser Arg Glu Pro Thr Ser Arg Pro Leu Trp Arg Glu Arg Ile Leu
                155                 160                 165
Gly Tyr Pro Ser Gln His Gly Val Cys Gln Ala Pro Ala Gly Gly
                170                 175                 180
Gly Arg Arg Val Gln Glu Gln Thr Ser Gln Leu His Gly Arg Arg
                185                 190                 195
Arg Gln Asp Pro Leu Pro Gly Gly Leu Asp Glu Asp Glu Gly Gln
                200                 205                 210
Ala Gly Leu Arg Ala Arg Gly Pro Leu Pro Gly Ser Pro Pro Glu
                215                 220                 225
Lys Lys Arg Phe His Leu Gly Ser Gln Thr Gly Cys Thr Gln Ser
                230                 235                 240
Gly Arg Gly Arg Glu Glu Asp Pro Gln Pro His Glu Asp Pro Pro
                245                 250                 255
Ala Gly Gln Gly Val Cys Gly His Arg Cys Arg Gln Lys Gly Ile
                260                 265                 270
```

```
<210> 4
<211> 225
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507098CD1

<400> 4
Met Arg Gly Glu Gly Gly Ala Gly Arg Cys Arg Gly Tyr Trp Glu
  1               5                  10                  15
Arg Ile Ser Arg Gly Arg Gly Val Val Pro Gly Arg Gly Ala Ala
                 20                  25                  30
Ala Gly Ser Ala Trp Pro Pro Gly Ala Met Ala Thr Leu Ser Phe
                 35                  40                  45
Val Phe Leu Leu Leu Gly Ala Val Ser Trp Pro Pro Ala Ser Ala
                 50                  55                  60
Ser Gly Gln Glu Phe Trp Pro Gly Gln Ser Ala Ala Asp Ile Leu
```

65 70 75
Ser Gly Ala Ala Ser Arg Arg Arg Tyr Leu Leu Tyr Asp Val Asn
80 85 90
Pro Pro Glu Gly Phe Asn Leu Arg Arg Asp Val Tyr Ile Arg Ile
95 100 105
Ala Ser Leu Leu Lys Thr Leu Leu Lys Thr Glu Glu Trp Val Leu
110 115 120
Val Leu Pro Pro Trp Gly Arg Leu Tyr His Trp Gln Ser Pro Asp
125 130 135
Ile His Gln Val Arg Ile Pro Trp Ser Glu Phe Phe Asp Leu Pro
140 145 150
Ser Leu Asn Lys Asn Ile Pro Val Ile Glu Tyr Glu Gln Phe Ile
155 160 165
Ala Glu Ser Gly Gly Ser Phe Ile Asp Gln Val Tyr Val Leu Gln
170 175 180
Ser Tyr Ala Glu Gly Trp Lys Glu Gly Thr Trp Glu Glu Lys Val
185 190 195
Asp Glu Arg Pro Cys Ile Asp Gln Leu Leu Tyr Ser Gln Asp Lys
200 205 210
His Glu Tyr Tyr Arg Ala Gln Pro Pro Ser Trp Arg Pro Cys Cys
215 220 225

<210> 5
<211> 283
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507099CD1

<400> 5
Met Arg Gly Glu Gly Gly Ala Gly Arg Cys Arg Gly Tyr Trp Glu
1 5 10 15
Arg Ile Ser Arg Gly Arg Gly Val Val Pro Gly Arg Gly Ala Ala
20 25 30
Ala Gly Ser Ala Trp Pro Pro Gly Ala Met Ala Thr Leu Ser Phe
35 40 45
Val Phe Leu Leu Leu Gly Ala Val Ser Trp Pro Pro Ala Ser Ala
50 55 60
Ser Gly Gln Glu Phe Trp Pro Gly Gln Ser Ala Ala Asp Ile Leu
65 70 75
Ser Gly Ala Ala Ser Arg Arg Arg Tyr Leu Leu Tyr Asp Val Asn
80 85 90
Pro Pro Glu Gly Phe Asn Leu Arg Arg Asp Val Tyr Ile Arg Ile
95 100 105
Ala Ser Leu Leu Lys Thr Leu Leu Lys Thr Glu Glu Trp Val Leu
110 115 120
Val Leu Pro Pro Trp Gly Arg Leu Tyr His Trp Gln Ser Pro Asp
125 130 135
Ile His Gln Val Arg Ile Pro Trp Ser Glu Phe Phe Asp Leu Pro
140 145 150
Ser Leu Asn Lys Asn Ile Pro Val Ile Glu Tyr Glu Gln Phe Ile
155 160 165
Ala Glu Ser Gly Gly Pro Phe Ile Asp Gln Val Tyr Val Leu Gln

```
                170                 175                 180
Ser Tyr Ala Glu Gly Trp Lys Glu Gly Thr Trp Glu Glu Lys Val
                185                 190                 195
Asp Glu Arg Pro Cys Ile Asp Gln Leu His Gly Arg Arg Arg Gln
                200                 205                 210
Asp Pro Leu Pro Gly Gly Leu Asp Glu Asp Glu Gly Gln Ala Gly
                215                 220                 225
Leu Arg Ala Arg Gly Pro Leu Pro Gly Ser Pro Pro Glu Lys Lys
                230                 235                 240
Arg Phe His Leu Gly Ser Gln Thr Gly Cys Thr Gln Ser Gly Arg
                245                 250                 255
Gly Arg Glu Glu Asp Pro Gln Pro His Glu Asp Pro Pro Ala Gly
                260                 265                 270
Gln Gly Val Cys Gly His Arg Cys Arg Gln Lys Gly Ile
                275                 280

<210> 6
<211> 294
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7501399CD1

<400> 6
Met Ser Arg Ser Ala Thr Leu Leu Leu Cys Leu Leu Gly Cys His
  1               5                  10                  15
Val Trp Lys Ala Val Thr Lys Thr Leu Arg Glu Pro Gly Ala Gly
                 20                  25                  30
Ala Gln Glu Val Thr Leu Lys Val His Ile Ser Asp Ala Ser Thr
                 35                  40                  45
His Gln Pro Val Ala Asp Ala Leu Ile Glu Ile Phe Thr Asn Gln
                 50                  55                  60
Ala Ser Ile Ala Ser Gly Thr Ser Gly Thr Asp Gly Val Ala Phe
                 65                  70                  75
Ile Lys Phe Gln Tyr Lys Leu Gly Ser Gln Leu Ile Val Thr Ala
                 80                  85                  90
Ser Lys His Ala Tyr Val Pro Asn Ser Ala Pro Trp Lys Pro Ile
                 95                 100                 105
Arg Leu Pro Val Phe Ser Ser Leu Ser Leu Gly Leu Leu Pro Glu
                110                 115                 120
Arg Ser Ala Thr Leu Met Val Tyr Glu Asp Val Val Gln Ile Val
                125                 130                 135
Ser Gly Phe Gln Gly Ala Arg Pro Gln Pro Arg Val His Phe Gln
                140                 145                 150
Arg Arg Ala Leu Arg Leu Pro Glu Asn Thr Ser Tyr Ser Asp Leu
                155                 160                 165
Thr Ala Phe Leu Thr Ala Ala Ser Ser Pro Ser Glu Val Asp Ser
                170                 175                 180
Phe Pro Tyr Leu Arg Gly Leu Asp Gly Asn Gly Thr Gly Asn Ser
                185                 190                 195
Thr Arg His Asp Leu Thr Pro Val Thr Ala Val Ser Val His Leu
                200                 205                 210
Leu Ser Ser Asn Gly Thr Pro Val Leu Val Asp Gly Pro Ile Tyr
                215                 220                 225
```

```
Val Thr Val Pro Leu Ala Thr Gln Ser Ser Leu Arg His Asn Ala
                230                 235                 240
Tyr Val Ala Ala Trp Arg Phe Asp Gln Lys Leu Gly Thr Trp Leu
                245                 250                 255
Lys Ser Gly Leu Gly Leu Val His Gln Glu Gly Ser Gln Leu Thr
                260                 265                 270
Trp Thr Tyr Ile Ala Pro Gln Leu Gly Tyr Trp Val Ala Ala Met
                275                 280                 285
Ser Pro Pro Ile Pro Gly Gly Ser Ala
                290

<210> 7
<211> 40
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504768CD1

<400> 7
Met Trp Val Pro Val Val Phe Leu Thr Leu Ser Val Thr Trp Ile
  1               5                  10                  15
Gly Gln Ser Ala Ala Val Phe Trp Cys Thr Pro Ser Gly Ser Ser
                 20                  25                  30
Gln Leu Pro Thr Ala Ser Gly Thr Lys Ala
                 35                  40

<210> 8
<211> 187
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500757CD1

<400> 8
Met Gly Ala Pro Leu Ala Val Ala Leu Gly Ala Leu His Tyr Leu
  1               5                  10                  15
Ala Leu Phe Leu Gln Leu Gly Gly Ala Thr Arg Pro Ala Gly His
                 20                  25                  30
Ala Pro Trp Asp Asn His Val Ser Gly His Ala Leu Phe Thr Glu
                 35                  40                  45
Thr Pro His Asp Met Thr Ala Arg Thr Gly Glu Asp Val Glu Met
                 50                  55                  60
Ala Cys Ser Phe Arg Gly Ser Gly Ser Pro Ser Tyr Ser Leu Glu
                 65                  70                  75
Ile Gln Trp Trp Tyr Val Arg Ser His Arg Asp Trp Thr Asp Lys
                 80                  85                  90
Gln Ala Trp Ala Ser Asn Gln Val Val Lys Val Val Gly Ser Asn
                 95                 100                 105
Ile Ser His Lys Leu Arg Leu Ser Arg Val Lys Pro Thr Asp Glu
                110                 115                 120
Gly Thr Tyr Glu Cys Arg Val Ile Asp Phe Ser Asp Gly Lys Ala
                125                 130                 135
```

```
Arg His His Lys Val Lys Ala Tyr Leu Arg Val Gln Pro Gly Glu
                140                 145                 150
Asn Ser Val Leu His Leu Pro Glu Ala Pro Pro Ala Ala Pro Ala
                155                 160                 165
Pro Pro Pro Pro Lys Pro Gly Lys Glu Leu Arg Lys Arg Ser Val
                170                 175                 180
Asp Gln Glu Ala Cys Ser Leu
                185
```

<210> 9
<211> 261
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1730616CD1

<400> 9

```
Met Lys Ile Thr Asn Gly Arg His Gly Asp Ser Ala Gly Ala Glu
  1               5                  10                  15
Gly Thr Met Glu Asn Phe Thr Ala Leu Phe Gly Ala Gln Ala Asp
                 20                  25                  30
Pro Pro Pro Pro Pro Thr Ala Leu Gly Phe Gly Pro Gly Lys Pro
                 35                  40                  45
Pro Pro Pro Pro Pro Pro Pro Ala Gly Gly Gly Pro Gly Thr Ala
                 50                  55                  60
Pro Pro Pro Thr Val Ala Thr Ala Pro Pro Gly Ala Asp Lys Ser
                 65                  70                  75
Gly Ala Gly Cys Gly Pro Phe Tyr Leu Met Arg Glu Leu Pro Gly
                 80                  85                  90
Ser Thr Glu Leu Thr Gly Ser Thr Asn Leu Ile Thr His Tyr Asn
                 95                 100                 105
Leu Glu Gln Ala Tyr Asn Lys Phe Cys Gly Lys Lys Val Lys Glu
                110                 115                 120
Lys Leu Ser Asn Phe Leu Pro Asp Leu Pro Gly Met Ile Asp Leu
                125                 130                 135
Pro Gly Ser His Asp Asn Ser Ser Leu Arg Ser Leu Ile Glu Lys
                140                 145                 150
Pro Pro Ile Leu Ser Ser Ser Phe Asn Pro Ile Thr Gly Thr Met
                155                 160                 165
Leu Ala Gly Phe Arg Leu His Thr Gly Pro Leu Pro Glu Gln Cys
                170                 175                 180
Arg Leu Met His Ile Gln Pro Pro Lys Lys Lys Asn Lys His Lys
                185                 190                 195
His Lys Gln Ser Arg Thr Gln Asp Pro Val Pro Pro Glu Thr Pro
                200                 205                 210
Ser Asp Ser Asp His Lys Lys Lys Lys Lys Lys Lys Glu Glu Asp
                215                 220                 225
Pro Glu Arg Lys Arg Lys Lys Lys Glu Lys Lys Lys Lys Lys Asn
                230                 235                 240
Arg His Ser Pro Asp His Pro Gly Met Gly Ser Ser Gln Ala Ser
                245                 250                 255
Ser Ser Ser Ser Leu Arg
                260
```

<210> 10
<211> 506
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 190404CD1

<400> 10
```
Met Ile Asn Ser Ile Phe Glu Ala Ile Leu Gln Ile Leu Ser His
 1               5                  10                  15
Pro Pro Ser Arg Arg Glu His Gly Tyr Asp Ala Val Val Leu Leu
                20                  25                  30
Ala Leu Leu Val Asn Tyr Arg Lys Tyr Glu Ser Val Asn Pro Tyr
                35                  40                  45
Ile Val Lys Leu Ser Ile Val Asp Asp Glu Ala Thr Leu Asn Gly
                50                  55                  60
Met Gly Leu Val Ile Ala Gln Ala Leu Ser Glu Tyr Asn Arg Gln
                65                  70                  75
Tyr Lys Asp Lys Glu Glu Glu His Gln Ser Gly Phe Phe Ser Ala
                80                  85                  90
Leu Thr Asn Met Val Gly Ser Met Phe Ile Ala Asp Ala His Glu
                95                 100                 105
Lys Ile Ser Val Gln Thr Asn Glu Ala Ile Leu Leu Ala Leu Tyr
               110                 115                 120
Glu Ala Val His Leu Asn Arg Asn Phe Ile Thr Val Leu Ala Gln
               125                 130                 135
Ser His Pro Glu Met Gly Leu Val Thr Thr Pro Val Ser Pro Ala
               140                 145                 150
Pro Thr Thr Pro Val Thr Pro Leu Gly Thr Thr Pro Pro Ser Ser
               155                 160                 165
Asp Val Ile Ser Ser Val Glu Leu Pro Leu Asp Ala Asp Val Gln
               170                 175                 180
Thr Ser Asn Leu Leu Ile Thr Phe Leu Lys Tyr Ser Ser Ile Val
               185                 190                 195
Met Gln Asp Thr Lys Asp Glu His Arg Leu His Ser Gly Lys Leu
               200                 205                 210
Cys Leu Ile Ile Leu Thr Cys Ile Ala Glu Asp Gln Tyr Ala Asn
               215                 220                 225
Ala Phe Leu His Asp Asp Asn Met Asn Phe Arg Val Asn Leu His
               230                 235                 240
Arg Met Pro Met Arg His Arg Lys Lys Ala Ala Asp Lys Asn Leu
               245                 250                 255
Pro Cys Arg Pro Leu Val Cys Ala Val Leu Asp Leu Met Val Glu
               260                 265                 270
Phe Ile Val Thr His Met Met Lys Glu Phe Pro Met Asp Leu Tyr
               275                 280                 285
Ile Arg Cys Ile Gln Val Val His Lys Leu Leu Cys Tyr Gln Lys
               290                 295                 300
Lys Cys Arg Val Arg Leu His Tyr Thr Trp Arg Glu Leu Trp Ser
               305                 310                 315
Ala Leu Ile Asn Leu Leu Lys Phe Leu Met Ser Asn Glu Thr Val
               320                 325                 330
Leu Leu Ala Lys His Asn Ile Phe Thr Leu Ala Leu Met Ile Val
               335                 340                 345
```

Asn Leu Phe Asn Met Phe Ile Thr Tyr Gly Asp Thr Phe Leu Pro
350 355 360
Thr Pro Ser Ser Tyr Asp Glu Leu Tyr Tyr Glu Ile Ile Arg Met
365 370 375
His Gln Ser Phe Asp Asn Leu Tyr Ser Met Val Leu Arg Leu Ser
380 385 390
Thr Asn Ala Gly Gln Trp Lys Glu Ala Ala Ser Lys Val Thr His
395 400 405
Ala Leu Val Asn Ile Arg Ala Ile Ile Asn His Phe Asn Pro Lys
410 415 420
Ile Glu Ser Tyr Ala Ala Val Asn His Ile Ser Gln Leu Ser Glu
425 430 435
Glu Gln Val Leu Glu Val Val Arg Ala Asn Tyr Asp Thr Leu Thr
440 445 450
Leu Lys Leu Gln Asp Gly Leu Asp Gln Tyr Glu Arg Tyr Ser Glu
455 460 465
Gln His Lys Glu Ala Ala Phe Phe Lys Glu Leu Val Arg Ser Ile
470 475 480
Ser Thr Asn Val Arg Arg Asn Leu Ala Phe His Thr Leu Ser Gln
485 490 495
Glu Val Leu Leu Lys Glu Phe Ser Thr Ile Ser
500 505

<210> 11
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500679CD1

<400> 11
Met Gly Glu Leu Ser Pro Leu Gln Arg Pro Leu Ala Thr Glu Gly
1 5 10 15
Thr Met Lys Ala Gln Gly Val Leu Leu Lys Leu Ala Leu Leu Ala
20 25 30
Leu Pro Leu Leu Leu Leu Leu Ser Thr Pro Pro Cys Ala Pro Gln
35 40 45
Val Ser Gly Ile Arg Gly Asp Gly Phe Pro Glu Glu Ile Gln Trp
50 55 60
Leu Ser Lys Phe Leu Pro Arg Leu Glu
65

<210> 12
<211> 121
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500687CD1

<400> 12
Met Ala Leu Pro Pro Gly Pro Ala Ala Leu Arg His Thr Leu Leu
1 5 10 15

Leu Leu Pro Ala Leu Leu Ser Ser Gly Pro Ser Arg His Pro Ser
20 25 30
Leu Ile Ser Ser Asp Ser Asn Asn Leu Lys Leu Asn Asn Val Arg
35 40 45
Leu Pro Arg Glu Asn Met Ser Leu Pro Ser Asn Leu Gln Leu Asn
50 55 60
Asp Leu Thr Pro Asp Ser Arg Ala Val Lys Pro Ala Asp Arg Gln
65 70 75
Met Ala Gln Asn Asn Ser Arg Pro Glu Leu Leu Asp Pro Glu Pro
80 85 90
Gly Gly Leu Leu Thr Ser Gln Gly Phe Ile Arg Leu Pro Val Leu
95 100 105
Gly Tyr Ile Tyr Arg Val Ser Ser Val Ser Ser Asp Glu Ile Trp
110 115 120
Leu

<210> 13
<211> 290
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500688CD1

<400> 13
Met Ala Leu Pro Pro Gly Pro Ala Ala Leu Arg His Thr Leu Leu
1 5 10 15
Leu Leu Pro Ala Leu Leu Ser Ser Gly Gly Pro Gly Thr Pro Arg
20 25 30
Leu Ala Trp Tyr Leu Asp Gly Gln Leu Gln Glu Ala Ser Thr Ser
35 40 45
Arg Leu Leu Ser Val Gly Gly Glu Ala Phe Ser Gly Gly Thr Ser
50 55 60
Thr Phe Thr Val Thr Ala His Arg Ala Gln His Glu Leu Asn Cys
65 70 75
Ser Leu Gln Asp Pro Arg Ser Gly Arg Ser Ala Asn Ala Ser Val
80 85 90
Ile Leu Asn Val Gln Phe Lys Pro Glu Ile Ala Gln Val Gly Ala
95 100 105
Lys Tyr Gln Glu Ala Gln Gly Pro Gly Leu Leu Val Val Leu Phe
110 115 120
Ala Leu Val Arg Ala Asn Pro Pro Ala Asn Val Thr Trp Ile Asp
125 130 135
Gln Asp Gly Pro Val Thr Val Asn Thr Ser Asp Phe Leu Val Leu
140 145 150
Asp Ala Gln Asn Tyr Pro Trp Leu Thr Asn His Thr Val Gln Leu
155 160 165
Gln Leu Arg Ser Leu Ala His Asn Leu Ser Val Val Ala Thr Asn
170 175 180
Asp Val Gly Val Thr Ser Ala Ser Leu Pro Ala Pro Gly Pro Ser
185 190 195
Arg His Pro Ser Leu Ile Ser Ser Asp Ser Asn Asn Leu Lys Leu
200 205 210
Asn Asn Val Arg Leu Pro Arg Glu Asn Met Ser Leu Pro Ser Asn

215 220 225
Leu Gln Leu Asn Asp Leu Thr Pro Asp Ser Arg Ala Val Lys Pro
230 235 240
Ala Asp Arg Gln Met Ala Gln Asn Asn Ser Arg Pro Glu Leu Leu
245 250 255
Asp Pro Glu Pro Gly Gly Leu Leu Thr Ser Gln Gly Phe Ile Arg
260 265 270
Leu Pro Val Leu Gly Tyr Ile Tyr Arg Val Ser Ser Val Ser Ser
275 280 285
Asp Glu Ile Trp Leu
290

<210> 14
<211> 27
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500697CD1

<400> 14
Met Ser Ala Leu Ser Leu Leu Ile Leu Gly Leu Leu Thr Ala Val
1 5 10 15
Pro Pro Ala Ser Cys Gln Gln Gly Ser Leu His Thr
20 25

<210> 15
<211> 500
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500709CD1

<400> 15
Met Leu Asp His Lys Asp Leu Glu Ala Glu Ile His Pro Leu Lys
1 5 10 15
Asn Glu Glu Arg Lys Ser Gln Glu Asn Leu Gly Asn Pro Ser Lys
20 25 30
Asn Glu Asp Asn Val Lys Ser Ala Pro Pro Gln Ser Arg Leu Ser
35 40 45
Arg Cys Arg Ala Ala Ala Phe Phe Leu Ser Leu Phe Leu Cys Leu
50 55 60
Phe Val Val Phe Val Val Ser Phe Val Ile Pro Cys Pro Asp Arg
65 70 75
Pro Ala Ser Gln Arg Met Trp Arg Ile Asp Tyr Ser Ala Ala Val
80 85 90
Ile Tyr Asp Phe Leu Ala Val Asp Asp Ile Asn Gly Asp Arg Ile
95 100 105
Gln Asp Val Leu Phe Leu Tyr Lys Asn Thr Asn Ser Ser Asn Asn
110 115 120
Phe Ser Arg Ser Cys Val Asp Glu Gly Phe Ser Ser Pro Cys Thr
125 130 135
Phe Ala Ala Ala Val Ser Gly Ala Asn Gly Ser Thr Leu Trp Glu

```
                140                 145                 150
Arg Pro Val Ala Gln Asp Val Ala Leu Val Glu Cys Ala Val Pro
                155                 160                 165
Gln Pro Arg Gly Ser Glu Ala Pro Ser Ala Cys Ile Leu Val Gly
                170                 175                 180
Arg Pro Ser Ser Phe Ile Ala Val Asn Leu Phe Thr Gly Glu Thr
                185                 190                 195
Leu Trp Asn His Ser Ser Ser Phe Ser Gly Asn Ala Ser Ile Leu
                200                 205                 210
Ser Pro Leu Leu Gln Val Pro Asp Val Asp Gly Asp Gly Ala Pro
                215                 220                 225
Asp Leu Leu Val Leu Thr Gln Glu Arg Glu Glu Val Ser Gly His
                230                 235                 240
Leu Tyr Ser Gly Ser Thr Gly His Gln Ile Gly Leu Arg Gly Ser
                245                 250                 255
Leu Gly Val Asp Gly Glu Ser Gly Phe Leu Leu His Val Thr Arg
                260                 265                 270
Thr Gly Ala His Tyr Ile Leu Phe Pro Cys Ala Ser Ser Leu Cys
                275                 280                 285
Gly Cys Ser Val Lys Gly Leu Tyr Glu Lys Val Thr Glu Ser Gly
                290                 295                 300
Gly Pro Phe Lys Ser Asp Pro His Trp Glu Ser Met Leu Asn Ala
                305                 310                 315
Thr Thr Arg Arg Met Leu Ser His Ser Ser Gly Ala Val Arg Tyr
                320                 325                 330
Leu Met His Val Pro Gly Asn Ala Gly Ala Asp Val Leu Leu Val
                335                 340                 345
Gly Ser Glu Ala Phe Val Leu Leu Asp Gly Gln Glu Leu Thr Pro
                350                 355                 360
Arg Trp Thr Pro Lys Ala Ala His Val Leu Arg Lys Pro Ile Phe
                365                 370                 375
Gly Arg Tyr Lys Pro Asp Thr Leu Ala Val Ala Val Glu Asn Gly
                380                 385                 390
Thr Gly Thr Asp Arg Gln Glu Thr Gly Glu Ala Arg His Ser Leu
                395                 400                 405
Tyr Met Phe His Pro Thr Leu Pro Arg Val Leu Leu Glu Leu Ala
                410                 415                 420
Asn Val Ser Thr His Ile Val Ala Phe Asp Ala Val Leu Phe Glu
                425                 430                 435
Pro Ser Arg His Ala Ala Tyr Ile Leu Leu Thr Gly Pro Ala Asp
                440                 445                 450
Ser Glu Ala Pro Gly Leu Val Ser Val Ile Lys His Lys Val Arg
                455                 460                 465
Asp Leu Val Pro Ser Ser Arg Val Val Arg Leu Gly Glu Gly Gly
                470                 475                 480
Pro Asp Ser Asp Gln Ala Ile Arg Asp Arg Phe Ser Arg Leu Arg
                485                 490                 495
Tyr Gln Ser Glu Ala
                500
```

<210> 16
<211> 543
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 7500711CD1

<400> 16

```
Met Leu Asp His Lys Asp Leu Glu Ala Glu Ile His Pro Leu Lys
 1               5                  10                  15
Asn Glu Glu Arg Lys Ser Gln Glu Asn Leu Gly Asn Pro Ser Lys
                20                  25                  30
Asn Glu Asp Asn Val Lys Ser Ala Pro Pro Gln Ser Arg Leu Ser
                35                  40                  45
Arg Cys Arg Ala Ala Ala Phe Phe Leu Ser Leu Phe Leu Cys Leu
                50                  55                  60
Phe Val Val Phe Val Val Ser Phe Val Ile Pro Cys Pro Asp Arg
                65                  70                  75
Pro Ala Ser Gln Arg Met Trp Arg Ile Asp Tyr Ser Ala Ala Val
                80                  85                  90
Ile Tyr Asp Phe Leu Ala Val Asp Asp Ile Asn Gly Asp Arg Ile
                95                 100                 105
Gln Asp Val Leu Phe Leu Tyr Lys Asn Thr Asn Ser Ser Asn Asn
               110                 115                 120
Phe Ser Arg Ser Cys Val Asp Glu Ala Ala Val Ser Gly Ala Asn
               125                 130                 135
Gly Ser Thr Leu Trp Glu Arg Pro Val Ala Gln Asp Val Ala Leu
               140                 145                 150
Val Glu Cys Ala Val Pro Gln Pro Arg Gly Ser Glu Ala Pro Ser
               155                 160                 165
Ala Cys Ile Leu Val Gly Arg Pro Ser Ser Phe Ile Ala Val Asn
               170                 175                 180
Leu Phe Thr Gly Glu Thr Leu Trp Asn His Ser Ser Ser Phe Ser
               185                 190                 195
Gly Asn Ala Ser Ile Leu Ser Pro Leu Leu Gln Val Pro Asp Val
               200                 205                 210
Asp Gly Asp Gly Ala Pro Asp Leu Leu Val Leu Thr Gln Glu Arg
               215                 220                 225
Glu Glu Val Ser Gly His Leu Tyr Ser Gly Ser Thr Gly His Gln
               230                 235                 240
Ile Gly Leu Arg Gly Ser Leu Gly Val Asp Gly Glu Ser Gly Phe
               245                 250                 255
Leu Leu His Val Thr Arg Thr Gly Ala His Tyr Ile Leu Phe Pro
               260                 265                 270
Cys Ala Ser Ser Leu Cys Gly Cys Ser Val Lys Gly Leu Tyr Glu
               275                 280                 285
Lys Val Thr Gly Ser Gly Gly Pro Phe Lys Ser Asp Pro His Trp
               290                 295                 300
Glu Ser Met Leu Asn Ala Thr Thr Arg Arg Met Leu Ser His Ser
               305                 310                 315
Ser Gly Ala Val Arg Tyr Leu Met His Val Pro Gly Asn Ala Gly
               320                 325                 330
Ala Asp Val Leu Leu Val Gly Ser Glu Ala Phe Val Leu Leu Asp
               335                 340                 345
Gly Gln Glu Leu Thr Pro Arg Trp Thr Pro Lys Ala Ala His Val
               350                 355                 360
Leu Arg Lys Pro Ile Phe Gly Arg Tyr Lys Pro Asp Thr Leu Ala
               365                 370                 375
Val Ala Val Glu Asn Gly Thr Gly Thr Asp Arg Gln Ile Leu Phe
               380                 385                 390
```

Leu Asp Leu Gly Thr Gly Ala Val Leu Cys Ser Leu Ala Leu Pro
395 400 405
Ser Leu Pro Gly Gly Pro Leu Ser Ala Ser Leu Pro Thr Ala Asp
410 415 420
His Arg Ser Ala Phe Phe Phe Trp Gly Leu His Glu Leu Gly Ser
425 430 435
Thr Ser Glu Thr Glu Thr Gly Glu Ala Arg His Ser Leu Tyr Met
440 445 450
Phe His Pro Thr Leu Pro Arg Val Leu Leu Glu Leu Ala Asn Val
455 460 465
Ser Thr His Ile Val Ala Phe Asp Ala Val Leu Phe Glu Pro Ser
470 475 480
Arg His Ala Ala Tyr Ile Leu Leu Thr Gly Pro Ala Asp Ser Glu
485 490 495
Ala Pro Gly Leu Val Ser Val Ile Lys His Lys Val Arg Asp Leu
500 505 510
Val Pro Ser Ser Arg Val Val Arg Leu Gly Glu Gly Gly Pro Asp
515 520 525
Ser Asp Gln Ala Ile Arg Asp Arg Phe Ser Arg Leu Arg Tyr Gln
530 535 540
Ser Glu Ala

<210> 17
<211> 72
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500723CD1

<400> 17
Met Gly Val Pro Thr Ala Pro Glu Ala Gly Ser Trp Arg Trp Gly
1 5 10 15
Ser Leu Leu Phe Ala Leu Phe Leu Ala Ala Ser Leu Gly Pro Ser
20 25 30
Pro Ala Pro Trp Arg Pro Pro Gly Cys Gln His Gln Pro Arg Pro
35 40 45
Gly Ser Ala Pro Arg Ala Gly Val Gly Leu Arg Pro Pro Trp Gln
50 55 60
Leu Leu His His His Ala Gln Pro Asp Pro Ala Gly
65 70

<210> 18
<211> 22
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500764CD1

<400> 18
Met Ser Trp Pro Arg Arg Leu Leu Leu Arg Tyr Leu Phe Pro Ala
1 5 10 15

Leu Leu Leu His Ala Val Lys
20

<210> 19
<211> 26
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500772CD1

<400> 19
Met Phe Leu Lys Ala Val Val Leu Thr Leu Ala Leu Val Ala Val
1 5 10 15
Ala Val Pro Ser Ser Arg Thr Asn Leu Glu Lys
20 25

<210> 20
<211> 27
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7501350CD1

<400> 20
Met Ile Pro Ala Val Val Leu Leu Leu Leu Leu Leu Val Glu Gln
1 5 10 15
Ala Asp Pro Ser Ala Lys Gly Ser Tyr Asn Gln Leu
20 25

<210> 21
<211> 253
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506396CD1

<400> 21
Met Pro Gln Met Arg Gln Thr Pro Thr Asp Lys Pro Leu Cys Pro
1 5 10 15
Ser Arg Thr His Lys Val Leu Pro Ile Leu Glu Ile Leu Tyr His
20 25 30
Val Glu Glu Arg Asn Ser His His Val Tyr Met Ala Leu Ile Ile
35 40 45
Leu Leu Ile Leu Thr Glu Asp Asp Gly Phe Asn Arg Ser Ile His
50 55 60
Glu Val Ile Leu Lys Asn Ile Thr Trp Tyr Ser Glu Arg Val Leu
65 70 75
Thr Glu Ile Ser Leu Gly Ser Leu Leu Ile Leu Val Val Ile Arg
80 85 90
Thr Ile Gln Tyr Asn Met Thr Arg Thr Arg Asp Lys Tyr Leu His

```
                95                  100                 105
Thr Asn Cys Leu Ala Ala Leu Ala Asn Met Ser Ala Gln Phe Arg
                110                 115                 120
Ser Leu His Gln Tyr Ala Ala Gln Arg Ile Ile Ser Leu Phe Ser
                125                 130                 135
Leu Leu Ser Lys Lys His Asn Lys Val Leu Glu Gln Ala Thr Gln
                140                 145                 150
Ser Leu Arg Gly Ser Leu Ser Ser Asn Asp Val Pro Leu Pro Asp
                155                 160                 165
Tyr Val Ile Ser Phe Phe Ser Ser Arg Leu Leu Gln Ala Gly Ala
                170                 175                 180
Glu Leu Ser Val Glu Arg Val Leu Glu Ile Ile Lys Gln Gly Val
                185                 190                 195
Val Ala Leu Pro Lys Asp Arg Leu Lys Lys Phe Pro Glu Leu Lys
                200                 205                 210
Phe Lys Tyr Val Glu Glu Glu Gln Pro Glu Glu Phe Phe Ile Pro
                215                 220                 225
Tyr Val Trp Ser Leu Val Tyr Asn Ser Ala Val Gly Leu Tyr Trp
                230                 235                 240
Asn Pro Gln Asp Ile Gln Leu Phe Thr Met Asp Ser Asp
                245                 250

<210> 22
<211> 511
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505917CD1

<400> 22
Met Ala Ala Glu Gln Asp Pro Glu Ala Arg Ala Ala Ala Arg Pro
  1               5                  10                  15
Leu Leu Thr Asp Leu Tyr Gln Ala Thr Met Ala Leu Gly Tyr Trp
                 20                  25                  30
Arg Ala Gly Arg Ala Arg Asp Ala Ala Glu Phe Glu Leu Phe Phe
                 35                  40                  45
Arg Arg Cys Pro Phe Gly Gly Ala Phe Ala Leu Ala Ala Gly Leu
                 50                  55                  60
Arg Asp Cys Val Arg Phe Leu Arg Ala Phe Arg Leu Arg Asp Ala
                 65                  70                  75
Asp Val Gln Phe Leu Ala Ser Val Leu Pro Pro Asp Thr Asp Pro
                 80                  85                  90
Ala Phe Phe Glu His Leu Arg Ala Leu Asp Cys Ser Glu Val Thr
                 95                 100                 105
Val Arg Ala Leu Pro Glu Gly Ser Leu Ala Phe Pro Gly Val Pro
                110                 115                 120
Leu Leu Gln Val Ser Gly Pro Leu Leu Val Val Gln Leu Leu Glu
                125                 130                 135
Thr Pro Leu Leu Cys Leu Val Ser Tyr Ala Ser Leu Val Ala Thr
                140                 145                 150
Asn Ala Ala Arg Leu Arg Leu Ile Ala Gly Pro Glu Lys Arg Leu
                155                 160                 165
Leu Glu Met Gly Leu Arg Arg Ala Gln Gly Pro Asp Gly Gly Leu
                170                 175                 180
```

Thr Ala Ser Thr Tyr Ser Tyr Leu Gly Gly Phe Asp Ser Ser Ser
185 190 195
Asn Val Leu Ala Gly Gln Leu Arg Gly Val Pro Val Ala Gly Thr
200 205 210
Leu Ala His Ser Phe Val Thr Ser Phe Ser Gly Ser Glu Val Pro
215 220 225
Pro Asp Pro Met Leu Ala Pro Ala Ala Gly Glu Gly Pro Gly Val
230 235 240
Asp Leu Ala Ala Lys Ala Gln Val Trp Leu Glu Gln Val Cys Ala
245 250 255
His Leu Gly Leu Gly Val Gln Glu Pro His Pro Gly Glu Arg Ala
260 265 270
Ala Phe Val Ala Tyr Ala Leu Ala Phe Pro Arg Ala Phe Gln Gly
275 280 285
Leu Leu Asp Thr Tyr Ser Val Trp Arg Ser Gly Leu Pro Asn Phe
290 295 300
Leu Ala Val Ala Leu Ala Leu Gly Glu Leu Gly Tyr Arg Ala Val
305 310 315
Gly Val Arg Leu Asp Ser Gly Asp Leu Leu Gln Gln Ala Gln Glu
320 325 330
Ile Arg Lys Val Phe Arg Ala Ala Ala Ala Gln Phe Gln Val Pro
335 340 345
Trp Leu Glu Ser Val Leu Ile Val Val Ser Asn Asn Ile Asp Glu
350 355 360
Glu Ala Leu Ala Arg Leu Ala Gln Glu Leu Val Ala Val Gly Gly
365 370 375
Gln Pro Arg Met Lys Leu Thr Glu Asp Pro Glu Lys Gln Thr Leu
380 385 390
Pro Gly Ser Lys Ala Ala Phe Arg Leu Leu Gly Ser Asp Gly Ser
395 400 405
Pro Leu Met Asp Met Leu Gln Leu Ala Glu Glu Pro Val Pro Gln
410 415 420
Ala Gly Gln Glu Leu Arg Val Trp Pro Pro Gly Ala Gln Glu Pro
425 430 435
Cys Thr Val Arg Pro Ala Gln Val Glu Pro Leu Leu Arg Leu Cys
440 445 450
Leu Gln Gln Gly Gln Leu Cys Glu Pro Leu Pro Ser Leu Ala Glu
455 460 465
Ser Arg Ala Leu Ala Gln Leu Ser Leu Ser Arg Leu Ser Pro Glu
470 475 480
His Arg Arg Leu Arg Ser Pro Ala Gln Tyr Gln Val Val Leu Ser
485 490 495
Glu Arg Leu Gln Ala Leu Val Asn Ser Leu Cys Ala Gly Gln Ser
500 505 510
Pro

<210> 23
<211> 127
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500701CD1

<400> 23
Met Gly His Arg Phe Leu Arg Gly Leu Leu Thr Leu Leu Leu Pro
1 5 10 15
Pro Pro Pro Leu Tyr Thr Arg His Arg Met Leu Gly Pro Glu Ser
20 25 30
Val Pro Pro Pro Lys Arg Ser Arg Ser Lys Leu Met Ala Pro Pro
35 40 45
Arg Ile Gly Thr His Asn Gly Thr Phe His Cys Asp Glu Ala Leu
50 55 60
Ala Cys Ala Leu Leu Arg Leu Leu Pro Glu Tyr Arg Asp Ala Glu
65 70 75
Ile Val Arg Thr Arg Asp Pro Glu Lys Leu Ala Ser Cys Asp Ile
80 85 90
Val Val Asp Val Gly Gly Glu Tyr Asp Pro Arg Arg His Arg Tyr
95 100 105
Asp His His Gln Arg Gln Gly Ser Ser Val Gln Trp Ile Trp Phe
110 115 120
Lys Arg Ser Phe Cys Arg Asp
125

<210> 24
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500702CD1

<400> 24
Met Gly His Arg Phe Leu Arg Gly Leu Leu Thr Leu Leu Leu Pro
1 5 10 15
Pro Pro Pro Leu Tyr Thr Arg His Arg Met Leu Gly Pro Glu Ser
20 25 30
Val Pro Pro Pro Lys Arg Ser Arg Ser Lys Leu Met Ala Pro Pro
35 40 45
Arg Ile Gly Thr His Asn Gly Thr Phe His Cys Asp Glu Ala Leu
50 55 60
Ala Cys Ala Leu Leu Arg Leu Leu Pro Glu Tyr Arg Asp Ala Glu
65 70 75
Ile Val Arg Thr Arg Asp Pro Glu Lys Leu Ala Ser Cys Asp Ile
80 85 90
Val Val Asp Val Gly Gly Glu Tyr Asp Pro Arg Arg His Arg Tyr
95 100 105
Asp His His Gln Arg Cys Met Arg Thr Leu Trp Arg Arg Trp Met
110 115 120
Leu Trp Thr Met Gly Ser Pro Ser Gly Gln Arg Gly Ser Leu Asp
125 130 135
Met His

<210> 25
<211> 207
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6044343CD1

<400> 25
Met Pro Gly Val Leu His Leu His Val Arg Ala Leu Pro Arg His
1 5 10 15
Arg Glu Pro Ala Gly Leu Leu Ala Gln Arg Gln Ile Val Pro Arg
20 25 30
Gly Gln Ser Leu Val Val Cys Val Ser Pro Phe Ser Pro Leu Trp
35 40 45
Gly Ser Glu Pro Pro Ala Leu Leu Leu Trp Val His Gln Ser Pro
50 55 60
Pro Arg Cys Pro His Ala Asp Ala His Arg Pro Ala Arg Gly Asn
65 70 75
Thr Asp Met Arg Thr Arg Arg Pro Thr His Arg Arg Thr Gln Ser
80 85 90
Cys Thr Trp Arg His Val Arg Thr Arg Arg Pro Thr Arg Arg Pro
95 100 105
Thr His Arg Cys Thr Arg Thr Cys Thr Trp Arg His Arg Cys Val
110 115 120
His Met Gln Ala His Thr Gln Val Ser Val Asp Leu Cys Glu Gln
125 130 135
Met Arg Ala Asp Leu His Val Glu Thr Gln Thr Cys Arg Arg Thr
140 145 150
Lys Arg Ser Leu Lys Ala Asp Thr Gln Glu Gly Ala Ser His Leu
155 160 165
Trp Asp Ser Arg Asp Ala Val Ala Asn Thr Ser Val Val Pro Ser
170 175 180
Ser Leu Leu Asp Lys Thr Glu Ala Cys Ser Gly His Thr Trp Glu
185 190 195
Ala Glu Pro Arg Asp Trp Ala Gln Glu Pro Phe Cys
200 205

<210> 26
<211> 1008
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503990CD1

<400> 26
Met Ala Ala Met Ala Pro Ala Leu Thr Asp Ala Ala Ala Glu Ala
1 5 10 15
His His Ile Arg Phe Lys Leu Ala Pro Pro Ser Ser Thr Leu Ser
20 25 30
Pro Gly Ser Ala Glu Asn Asn Gly Asn Ala Asn Ile Leu Ile Ala
35 40 45
Ala Asn Gly Thr Lys Arg Lys Ala Ile Ala Ala Glu Asp Pro Ser
50 55 60
Leu Asp Phe Arg Asn Asn Pro Thr Lys Glu Asp Leu Gly Lys Leu
65 70 75
Gln Pro Leu Val Ala Ser Tyr Leu Cys Ser Asp Val Thr Ser Val
80 85 90

```
Pro Ser Lys Glu Ser Leu Lys Leu Gln Gly Val Phe Ser Lys Gln
                95                  100                 105
Thr Val Leu Lys Ser His Pro Leu Leu Ser Gln Ser Tyr Glu Leu
                110                 115                 120
Arg Ala Glu Leu Leu Gly Arg Gln Pro Val Leu Glu Phe Ser Leu
                125                 130                 135
Glu Asn Leu Arg Thr Met Asn Thr Ser Gly Gln Thr Ala Leu Pro
                140                 145                 150
Gln Ala Pro Val Asn Gly Leu Ala Lys Lys Leu Thr Lys Ser Ser
                155                 160                 165
Thr His Ser Asp His Asp Asn Ser Thr Ser Leu Asn Gly Gly Lys
                170                 175                 180
Arg Ala Leu Thr Ser Ser Ala Leu His Gly Gly Glu Met Gly Gly
                185                 190                 195
Ser Glu Ser Gly Asp Leu Lys Gly Gly Met Thr Asn Cys Thr Leu
                200                 205                 210
Pro His Arg Ser Leu Asp Val Glu His Thr Ile Leu Tyr Ser Asn
                215                 220                 225
Asn Ser Thr Ala Asn Lys Ser Ser Val Asn Ser Met Glu Gln Pro
                230                 235                 240
Ala Leu Gln Gly Ser Ser Arg Leu Ser Pro Gly Thr Asp Ser Ser
                245                 250                 255
Ser Asn Leu Gly Gly Val Lys Leu Glu Gly Lys Lys Ser Pro Leu
                260                 265                 270
Ser Ser Ile Leu Phe Ser Ala Leu Asp Ser Asp Thr Arg Ile Thr
                275                 280                 285
Ala Leu Leu Arg Arg Gln Ala Asp Ile Glu Ser Arg Ala Arg Arg
                290                 295                 300
Leu Gln Lys Arg Leu Gln Val Val Gln Ala Lys Gln Val Glu Arg
                305                 310                 315
His Ile Gln His Gln Leu Gly Gly Phe Leu Glu Lys Thr Leu Ser
                320                 325                 330
Lys Leu Pro Asn Leu Glu Ser Leu Arg Pro Arg Ser Gln Leu Met
                335                 340                 345
Leu Thr Arg Lys Ala Glu Ala Ala Leu Arg Lys Ala Ala Ser Glu
                350                 355                 360
Thr Thr Thr Ser Glu Gly Leu Ser Asn Phe Leu Lys Ser Asn Ser
                365                 370                 375
Ile Ser Glu Glu Leu Glu Arg Phe Thr Ala Ser Gly Ile Ala Asn
                380                 385                 390
Leu Arg Cys Ser Glu Gln Ala Phe Asp Ser Asp Val Thr Asp Ser
                395                 400                 405
Ser Ser Gly Gly Glu Ser Asp Ile Glu Glu Glu Glu Leu Thr Arg
                410                 415                 420
Ala Asp Pro Glu Gln Arg His Val Pro Leu Arg Arg Arg Ser Glu
                425                 430                 435
Trp Lys Trp Ala Ala Asp Arg Ala Ala Ile Val Ser Arg Trp Asn
                440                 445                 450
Trp Leu Gln Ala His Val Ser Asp Leu Glu Tyr Arg Ile Arg Gln
                455                 460                 465
Gln Thr Asp Ile Tyr Lys Gln Ile Arg Ala Asn Lys Ile Glu Ser
                470                 475                 480
Val Ser Gln Pro Leu Glu Asn His Gly Ala Pro Ile Ile Gly His
                485                 490                 495
Ile Ser Glu Ser Leu Ser Thr Lys Ser Cys Gly Ala Leu Arg Pro
                500                 505                 510
```

Val Asn Gly Val Ile Asn Thr Leu Gln Pro Val Leu Ala Asp His
515 520 525
Ile Pro Gly Asp Ser Ser Asp Ala Glu Glu Gln Leu His Lys Lys
530 535 540
Gln Arg Leu Asn Leu Val Ser Ser Ser Ser Asp Gly Thr Cys Val
545 550 555
Ala Ala Arg Thr Arg Pro Val Leu Ser Cys Lys Lys Arg Arg Leu
560 565 570
Val Arg Pro Asn Ser Ile Val Pro Leu Ser Lys Lys Val His Arg
575 580 585
Asn Ser Thr Ile Arg Pro Gly Cys Asp Val Asn Pro Ser Cys Ala
590 595 600
Leu Cys Gly Ser Gly Ser Ile Asn Thr Met Pro Pro Glu Ile His
605 610 615
Tyr Glu Ala Pro Leu Leu Glu Arg Leu Ser Gln Leu Asp Ser Cys
620 625 630
Val His Pro Val Leu Ala Phe Pro Asp Asp Val Pro Thr Ser Leu
635 640 645
His Phe Gln Ser Met Leu Lys Ser Gln Trp Gln Asn Lys Pro Phe
650 655 660
Asp Lys Ile Lys Pro Pro Lys Lys Leu Ser Leu Lys His Arg Ala
665 670 675
Pro Met Pro Gly Ser Leu Pro Asp Ser Ala Arg Lys Asp Arg His
680 685 690
Lys Leu Val Ser Ser Phe Leu Thr Thr Ala Met Leu Lys His His
695 700 705
Thr Asp Met Ser Ser Ser Ser Tyr Leu Ala Ala Thr His His Pro
710 715 720
Pro His Ser Pro Leu Val Arg Gln Leu Ser Thr Ser Ser Asp Ser
725 730 735
Pro Ala Pro Ala Ser Ser Ser Ser Gln Val Thr Ala Ser Thr Ser
740 745 750
Gln Gln Pro Val Arg Arg Arg Arg Gly Glu Ser Ser Phe Asp Ile
755 760 765
Asn Asn Ile Val Ile Pro Met Ser Val Ala Ala Thr Thr Arg Val
770 775 780
Glu Lys Leu Gln Tyr Lys Glu Ile Leu Thr Pro Ser Trp Arg Glu
785 790 795
Val Asp Leu Gln Ser Leu Lys Gly Ser Pro Asp Glu Glu Asn Glu
800 805 810
Glu Ile Glu Asp Leu Ser Asp Ala Ala Phe Ala Ala Leu His Ala
815 820 825
Lys Cys Glu Glu Met Glu Arg Ala Arg Trp Leu Trp Thr Thr Ser
830 835 840
Val Pro Pro Gln Arg Arg Gly Ser Arg Ser Tyr Arg Ser Ser Asp
845 850 855
Gly Arg Thr Thr Pro Gln Leu Gly Ser Ala Asn Pro Ser Thr Pro
860 865 870
Gln Pro Ala Ser Pro Asp Val Ser Ser Ser His Ser Leu Ser Glu
875 880 885
Tyr Ser His Gly Gln Ser Pro Arg Ser Pro Ile Ser Pro Glu Leu
890 895 900
His Ser Ala Pro Leu Thr Pro Val Ala Arg Asp Thr Leu Arg His
905 910 915
Leu Ala Ser Glu Asp Thr Arg Cys Ser Thr Pro Glu Leu Gly Leu
920 925 930

```
Asp Glu Gln Ser Val Gln Pro Trp Glu Arg Arg Thr Phe Pro Leu
                935                 940                 945
Ala His Ser Pro Gln Ala Glu Cys Glu Asp Gln Leu Asp Ala Gln
                950                 955                 960
Glu Arg Ala Ala Arg Cys Thr Arg Arg Thr Ser Gly Ser Lys Thr
                965                 970                 975
Gly Arg Glu Thr Glu Ala Ala Pro Thr Ser Pro Pro Ile Val Pro
                980                 985                 990
Leu Lys Ser Arg His Leu Val Ala Ala Ala Thr Ala Gln Arg Pro
                995                1000                1005
Thr His Arg


<210> 27
<211> 70
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504655CD1

<400> 27
Met Val Gly Pro Ala Pro Arg Arg Arg Leu Arg Pro Leu Ala Ala
  1               5                  10                  15
Leu Ala Leu Val Leu Ala Leu Ala Pro Gly Leu Pro Thr Ala Arg
                 20                  25                  30
Ala Gly Gln Thr Pro Arg Pro Ala Glu Arg Gly Pro Pro Val Arg
                 35                  40                  45
Leu Phe Thr Glu Glu Glu Leu Ala Arg Tyr Gly Gly Glu Glu Ser
                 50                  55                  60
Phe Met Asp Glu Glu Pro Pro Thr Met Pro
                 65                  70

<210> 28
<211> 142
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504690CD1

<400> 28
Met Leu Leu Trp Pro Leu Leu Leu Leu Leu Leu Leu Leu Pro Thr
  1               5                  10                  15
Leu Ala Leu Leu Arg Gln Gln Arg Ser Gln Asp Ala Arg Leu Ser
                 20                  25                  30
Trp Leu Ala Gly Leu Gln His Arg Val Ala Trp Gly Ala Leu Val
                 35                  40                  45
Trp Ala Ala Thr Trp Gln Arg Arg Arg Leu Glu Gln Ser Thr Leu
                 50                  55                  60
His Val His Gln Ser Gln Gln Gln Ala Leu Arg Trp Cys Leu Gln
                 65                  70                  75
Gly Ala Gln Arg Pro His Cys Ser Leu Arg Arg Ser Thr Asp Ile
                 80                  85                  90
```

```
Ser Thr Phe Arg Asn His Leu Pro Leu Thr Lys Ala Ser Gln Thr
                95                  100                 105
Gln Gln Glu Asp Ser Gly Arg Gly Pro Ser Cys Trp Thr Met Ala
                110                 115                 120
Val Trp Arg Ala Ala Phe Trp Ile Pro Leu Arg Ala Leu Leu Pro
                125                 130                 135
Thr Thr Arg Cys Leu Trp Arg
                140
```

<210> 29
<211> 43
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504720CD1

<400> 29

```
Met Ala Ser Arg Leu Thr Leu Leu Thr Leu Leu Leu Leu Leu Leu
  1             5                   10                  15
Ala Gly Gly Leu Gly Arg Thr Pro Lys Gln Thr Trp Arg Ala Ser
                20                  25                  30
Ser Leu Thr Pro Arg Thr Ser Pro Val Ser Thr Arg Pro
                35                  40
```

<210> 30
<211> 64
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504722CD1

<400> 30

```
Met Ala Ser Gly Ser Asn Trp Leu Ser Gly Val Asn Val Val Leu
  1             5                   10                  15
Val Met Ala Tyr Gly Ser Leu Val Phe Val Leu Leu Phe Ile Phe
                20                  25                  30
Val Lys Arg Gln Ile Met Arg Phe Ala Met Lys Ser Arg Arg Gly
                35                  40                  45
Pro His Val Pro Val Gly His Asn Ala Pro Lys Val Ala Thr Thr
                50                  55                  60
Ile Cys Ile Gly
```

<210> 31
<211> 23
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504733CD1

<400> 31
Met Lys Leu Leu Ala Ala Thr Val Leu Leu Leu Thr Ile Cys Ser
1 5 10 15
Leu Glu Val Leu Pro Asp Arg Asp
20

<210> 32
<211> 349
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507100CD1

<400> 32
Met Arg Gly Glu Gly Gly Ala Gly Arg Cys Arg Gly Tyr Trp Glu
1 5 10 15
Arg Ile Ser Arg Gly Arg Gly Val Val Pro Gly Arg Gly Ala Ala
20 25 30
Ala Gly Ser Ala Trp Pro Pro Gly Ala Met Ala Thr Leu Ser Phe
35 40 45
Val Phe Leu Leu Leu Gly Ala Val Ser Trp Pro Pro Ala Ser Ala
50 55 60
Ser Gly Gln Glu Phe Trp Pro Gly Gln Ser Ala Ala Asp Ile Leu
65 70 75
Ser Gly Ala Ala Ser Arg Arg Arg Tyr Leu Leu Tyr Asp Val Asn
80 85 90
Pro Pro Glu Gly Phe Asn Leu Arg Arg Asp Val Tyr Ile Arg Ile
95 100 105
Ala Ser Leu Leu Lys Thr Leu Leu Lys Thr Glu Glu Trp Val Leu
110 115 120
Val Leu Pro Pro Trp Gly Arg Leu Tyr His Trp Gln Ser Pro Asp
125 130 135
Ile His Gln Val Arg Ile Pro Trp Ser Glu Phe Phe Asp Leu Pro
140 145 150
Ser Leu Asn Lys Asn Ile Pro Val Ile Glu Tyr Glu Gln Phe Ile
155 160 165
Ala Glu Ser Gly Gly Pro Phe Ile Asp Gln Val Tyr Val Leu Gln
170 175 180
Ser Tyr Ala Glu Gly Trp Lys Glu Gly Thr Trp Glu Glu Lys Val
185 190 195
Asp Glu Arg Pro Cys Ile Asp Gln Leu Leu Tyr Ser Gln Asp Lys
200 205 210
His Glu Tyr Tyr Arg Cys Leu Leu Arg Leu Leu Pro Leu Pro Gln
215 220 225
Gly Ser Ala Ser Ile Val Ala Pro Leu Leu Leu Arg Asn Thr Ser
230 235 240
Ala Arg Ser Val Met Leu Asp Arg Ala Glu Asn Leu Leu His Asp
245 250 255
His Tyr Gly Gly Lys Glu Tyr Trp Asp Thr Arg Arg Ser Met Val
260 265 270
Phe Ala Arg His Leu Arg Glu Val Gly Asp Glu Phe Arg Ser Arg
275 280 285
His Leu Asn Ser Thr Asp Asp Ala Asp Arg Ile Pro Phe Gln Glu
290 295 300

Asp Trp Met Lys Met Lys Val Lys Leu Gly Ser Ala Leu Gly Gly
305 310 315
Pro Tyr Leu Gly Val His Leu Arg Arg Lys Asp Phe Ile Trp Gly
320 325 330
His Arg Gln Asp Val Pro Ser Leu Glu Gly Ala Gly Gly Thr Gly
335 340 345
Arg Pro Leu Ala

<210> 33
<211> 65
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503330CD1

<400> 33
Met Phe Val Ala Ser Glu Arg Lys Met Arg Ala His Gln Val Leu
1 5 10 15
Thr Phe Leu Leu Leu Phe Val Ile Thr Ser Val Ala Ser Glu His
20 25 30
Ala Gln Leu Leu Arg His Val Ala Ala Gln Asp Ala Gly Asp Gly
35 40 45
Leu Arg Gly Gly Arg Ala Ala Ala Ala Gly Leu Tyr Gly Glu Gln
50 55 60
Gly Leu Leu His Gly
65

<210> 34
<211> 64
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504519CD1

<400> 34
Met Asp His Cys Gly Ala Leu Phe Leu Cys Leu Cys Leu Leu Thr
1 5 10 15
Leu Gln Asn Ala Thr Thr Glu Thr Trp Glu Glu Leu Leu Ser Tyr
20 25 30
Met Glu Asn Met Gln Val Ser Arg Gly Arg Ser Ser Val Phe Ser
35 40 45
Ser Arg Gln Glu Val Ser Met Pro Gly Val Ser Thr Pro Cys Ser
50 55 60
Ser Pro Pro Ser

<210> 35
<211> 94
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504705CD1

<400> 35
Met Lys Ser Leu Ile Leu Leu Ala Ile Leu Ala Ala Leu Ala Val
1 5 10 15
Val Thr Leu Cys Tyr Glu Ser His Glu Ser Met Glu Ser Tyr Glu
20 25 30
Leu Arg Ser Glu Asn Ala Leu Ser Leu Ser Thr Ser Ser Ile Gly
35 40 45
Lys Pro Val Met Thr Thr Asp Phe Ala Asn Ala Thr Pro Trp Phe
50 55 60
Met Asp Thr Met Leu Pro Ile Ile Ala Thr Ser Gly Ser Ala Glu
65 70 75
Gly Pro Asn Glu Thr Glu Gly Arg Lys Lys Ile Ser Phe Phe Leu
80 85 90
Glu Ala Gly Thr

<210> 36
<211> 52
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504738CD1

<400> 36
Met Lys Leu Cys Val Thr Val Leu Ser Leu Leu Met Leu Val Ala
1 5 10 15
Ala Phe Cys Ser Pro Ala Leu Ser Ala Pro Asn Ser Lys Pro Lys
20 25 30
Glu Ala Ser Lys Ser Val Leu Ile Pro Val Asn Pro Gly Ser Arg
35 40 45
Ser Thr Cys Met Thr Trp Asn
50

<210> 37
<211> 33
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510280CD1

<400> 37
Met Asp Asn Val Gln Pro Lys Ile Lys His Arg Pro Phe Cys Phe
1 5 10 15
Ser Val Lys Gly His Val Lys Met Leu Arg Leu His Cys Leu Val
20 25 30
Val Phe Met

<210> 38
<211> 242
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503700CD1

<400> 38
Met Ser Pro Trp Ser Trp Phe Leu Leu Gln Thr Leu Cys Leu Leu
1 5 10 15
Pro Thr Gly Ala Ala Ser Arg Arg Gly Ala Pro Gly Thr Ala Asn
20 25 30
Cys Glu Leu Lys Pro Gln Gln Ser Glu Leu Asn Ser Phe Leu Trp
35 40 45
Thr Ile Lys Arg Asp Pro Pro Ser Tyr Phe Phe Gly Thr Ile His
50 55 60
Val Pro Tyr Thr Arg Val Trp Asp Phe Ile Pro Asp Asn Ser Lys
65 70 75
Glu Ala Phe Leu Gln Ser Ser Ile Val Tyr Phe Glu Leu Asp Leu
80 85 90
Thr Asp Pro Tyr Thr Ile Ser Ala Leu Thr Ser Cys Gln Met Leu
95 100 105
Pro Gln Gly Glu Asn Leu Gln Asp Val Leu Pro Arg Asp Ile Tyr
110 115 120
Cys Arg Leu Lys Arg His Leu Glu Tyr Val Lys Leu Met Met Pro
125 130 135
Leu Trp Met Thr Pro Asp Gln Arg Gly Lys Gly Leu Tyr Ala Asp
140 145 150
Tyr Leu Phe Asn Ala Ile Ala Gly Asn Trp Glu Arg Lys Arg Pro
155 160 165
Val Trp Val Met Leu Met Val Asn Ser Leu Thr Glu Val Asp Ile
170 175 180
Lys Ser Arg Gly Val Pro Val Leu Asp Leu Phe Leu Ala Gln Glu
185 190 195
Ala Glu Arg Leu Arg Lys Gln Thr Gly Ala Val Glu Lys Val Glu
200 205 210
Glu Gln Cys His Pro Leu Asn Gly Leu Asn Phe Ser Gln Val Ile
215 220 225
Ser Trp Ala Thr Thr Gln Cys Trp Met Phe Cys Gly Val Lys Ala
230 235 240
Met Arg

<210> 39
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504685CD1

<400> 39
Met Ala Ser Thr Val Val Ala Val Gly Leu Thr Ile Ala Ala Ala

1 5 10 15
Gly Phe Ala Gly Arg Tyr Val Leu Gln Ala Met Lys His Met Glu
20 25 30
Pro Gln Val Lys Gln Val Phe Gln Ser Leu Pro Lys Ser Pro Tyr
35 40 45
Cys Gln

<210> 40
<211> 53
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506844CD1

<400> 40
Met Ser Ser Phe Gly Tyr Arg Thr Leu Thr Val Ala Leu Phe Thr
1 5 10 15
Leu Ile Cys Cys Pro Gly Val Asn Glu Phe Gln Arg Gln Arg Val
20 25 30
Pro Ala Ser Lys Ala Gly His Pro Asp Thr Ala Thr His Phe Gly
35 40 45
Gly Cys Gly Gln Val Leu His His
50

<210> 41
<211> 95
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510259CD1

<400> 41
Met Met Tyr Ala Pro Val Glu Phe Ser Glu Ala Glu Phe Ser Arg
1 5 10 15
Ala Glu Tyr Gln Arg Lys Gln Gln Phe Trp Asp Ser Val Arg Leu
20 25 30
Ala Leu Phe Thr Leu Ala Ile Val Ala Ile Ile Gly Ile Ala Ile
35 40 45
Gly Ile Val Thr His Phe Val Val Glu Asp Asp Lys Ser Phe Tyr
50 55 60
Tyr Leu Ala Ser Phe Lys Val Thr Asn Ile Lys Tyr Lys Glu Asn
65 70 75
Tyr Gly Ile Arg Ser Ser Arg Glu Phe Ile Glu Arg Ser His Gln
80 85 90
Ile Glu Arg Met Arg
95

<210> 42
<211> 57
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510444CD1

<400> 42
Met Val Gly Gln Arg Val Leu Leu Leu Val Ala Phe Leu Leu Ser
1 5 10 15
Gly Val Leu Leu Ser Glu Ala Ala Lys Ile Leu Thr Ile Ser Thr
20 25 30
Leu Gly Glu Cys Leu Ala Gly Glu Phe Pro Asp Arg Arg Val Pro
35 40 45
Asp Pro Arg Thr Ala Arg Ala Pro Ala Asn Gly Asp
50 55

<210> 43
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510494CD1

<400> 43
Met Pro Gly Arg Ser Cys Val Ala Leu Val Leu Leu Ala Ala Ala
1 5 10 15
Val Ser Cys Ala Val Ala Gln His Ala Pro Pro Val Ser Glu Leu
20 25 30
Glu Pro Arg Arg Arg Glu Gly Arg Ala Val Asp Arg Gly Leu Gln
35 40 45
Lys Ile Asn Leu Ser Ser Phe Arg Thr Asn Val Gln Arg Cys Ser
50 55 60
Ser Met Val His His Lys Ser
65

<210> 44
<211> 311
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6486485CD1

<400> 44
Met Ala Phe Leu Val Ala His Pro Met Gln Phe Val Tyr Leu Leu
1 5 10 15
Thr Thr Leu Cys Val Phe Asn Met Val Phe Ala Lys Leu Gly Phe
20 25 30
Ser Glu Thr Val Phe Ser Gln Arg Leu Ser Phe Thr Val Leu Ser
35 40 45
Ala Val Gly Tyr Phe Gln Trp Gln Lys Arg Pro His Leu Leu Pro
50 55 60
Val Gly Pro Leu Gly Arg Ser Met Arg Trp Cys Leu Leu Leu Ile
65 70 75
Trp Ala Gln Gly Leu Arg Gln Ala Pro Leu Ala Ser Gly Met Met

```
                80                      85                      90
Thr Gly Thr Ile Glu Thr Thr Gly Asn Ile Ser Ala Glu Lys Gly
                95                      100                     105
Gly Ser Ile Ile Leu Gln Cys His Leu Ser Ser Thr Thr Ala Gln
                110                     115                     120
Val Thr Gln Val Asn Trp Glu Gln Gln Asp Gln Leu Leu Ala Ile
                125                     130                     135
Cys Asn Ala Asp Leu Gly Trp His Ile Ser Pro Ser Phe Lys Asp
                140                     145                     150
Arg Val Ala Pro Gly Pro Gly Leu Gly Leu Thr Leu Gln Ser Leu
                155                     160                     165
Thr Val Asn Asp Thr Gly Glu Tyr Phe Cys Ile Tyr His Thr Tyr
                170                     175                     180
Pro Asp Gly Thr Tyr Thr Gly Arg Ile Phe Leu Glu Val Leu Glu
                185                     190                     195
Ser Ser Val Ala Glu His Gly Ala Arg Phe Gln Ile Pro Leu Leu
                200                     205                     210
Gly Ala Met Ala Ala Thr Leu Val Val Ile Cys Thr Ala Val Ile
                215                     220                     225
Val Val Val Ala Leu Thr Arg Lys Lys Lys Ala Leu Arg Ile His
                230                     235                     240
Ser Val Glu Gly Asp Leu Arg Arg Lys Ser Ala Gly Gln Glu Glu
                245                     250                     255
Trp Ser Pro Ser Ala Pro Ser Pro Pro Gly Ser Cys Val Gln Ala
                260                     265                     270
Glu Ala Ala Pro Ala Gly Leu Cys Gly Glu Gln Arg Gly Glu Asp
                275                     280                     285
Cys Ala Glu Leu His Asp Tyr Phe Asn Val Leu Ser Tyr Arg Ser
                290                     295                     300
Leu Gly Asn Cys Ser Phe Phe Thr Glu Thr Gly
                305                     310
```

<210> 45
<211> 1097
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503772CD1

<400> 45

```
Met Gly Ala Ser Arg Asp Arg Gly Leu Ala Ala Leu Trp Cys Leu
 1              5                       10                      15
Gly Leu Leu Gly Gly Leu Ala Arg Val Ala Gly Thr His Tyr Arg
                20                      25                      30
Tyr Leu Trp Arg Gly Cys Tyr Pro Cys His Leu Gly Gln Ala Gly
                35                      40                      45
Tyr Pro Val Ser Ala Gly Asp Gln Arg Pro Asp Val Asp Glu Cys
                50                      55                      60
Arg Thr His Asn Gly Gly Cys Gln His Arg Cys Val Asn Thr Pro
                65                      70                      75
Gly Ser Tyr Leu Cys Glu Cys Lys Pro Gly Phe Arg Leu His Thr
                80                      85                      90
Asp Ser Arg Thr Cys Leu Ala Ile Asn Ser Cys Ala Leu Gly Asn
                95                      100                     105
```

Gly Gly Cys Gln His His Cys Val Gln Leu Thr Ile Thr Arg His
110 115 120
Arg Cys Gln Cys Arg Pro Gly Phe Gln Leu Gln Glu Asp Gly Arg
125 130 135
His Cys Val Arg Arg Ser Pro Cys Ala Asn Arg Asn Gly Ser Cys
140 145 150
Met His Arg Cys Gln Val Val Arg Gly Leu Ala Arg Cys Glu Cys
155 160 165
His Val Gly Tyr Gln Leu Ala Ala Asp Gly Lys Ala Cys Glu Asp
170 175 180
Val Asp Glu Cys Ala Ala Gly Leu Ala Gln Cys Ala His Gly Cys
185 190 195
Leu Asn Thr Gln Gly Ser Phe Lys Cys Val Cys His Ala Gly Tyr
200 205 210
Glu Leu Gly Ala Asp Gly Arg Gln Cys Tyr Arg Ile Glu Met Glu
215 220 225
Ile Val Asn Ser Cys Glu Ala Asn Asn Gly Gly Cys Ser His Gly
230 235 240
Cys Ser His Thr Ser Ala Gly Pro Leu Cys Thr Cys Pro Arg Gly
245 250 255
Tyr Glu Leu Asp Thr Asp Gln Arg Thr Cys Ile Asp Val Asp Asp
260 265 270
Cys Ala Asp Ser Pro Cys Cys Gln Gln Val Cys Thr Asn Asn Pro
275 280 285
Gly Gly Tyr Glu Cys Gly Cys Tyr Ala Gly Tyr Arg Leu Ser Ala
290 295 300
Asp Gly Cys Gly Cys Glu Asp Val Asp Glu Cys Ala Ser Ser Arg
305 310 315
Gly Gly Cys Glu His His Cys Thr Asn Leu Ala Gly Ser Phe Gln
320 325 330
Cys Ser Cys Glu Ala Gly Tyr Arg Leu His Glu Asp Arg Arg Gly
335 340 345
Cys Ser Pro Leu Glu Glu Pro Met Val Asp Leu Asp Gly Glu Leu
350 355 360
Pro Phe Val Arg Pro Leu Pro His Ile Ala Val Leu Gln Asp Glu
365 370 375
Leu Pro Gln Leu Phe Gln Asp Asp Asp Val Gly Ala Asp Glu Glu
380 385 390
Glu Ala Glu Leu Arg Gly Glu His Thr Leu Thr Glu Lys Phe Val
395 400 405
Cys Leu Asp Asp Ser Phe Gly His Asp Cys Ser Leu Thr Cys Asp
410 415 420
Asp Cys Arg Asn Gly Gly Thr Cys Leu Leu Gly Leu Asp Gly Cys
425 430 435
Asp Cys Pro Glu Gly Trp Thr Gly Leu Ile Cys Asn Glu Thr Cys
440 445 450
Pro Pro Asp Thr Phe Gly Lys Asn Cys Ser Phe Ser Cys Ser Cys
455 460 465
Gln Asn Gly Gly Thr Cys Asp Ser Val Thr Gly Ala Cys Arg Cys
470 475 480
Pro Pro Gly Val Ser Gly Thr Asn Cys Glu Asp Gly Cys Pro Lys
485 490 495
Gly Tyr Tyr Gly Lys His Cys Arg Lys Lys Cys Asn Cys Ala Asn
500 505 510
Arg Gly Arg Cys His Arg Leu Tyr Gly Ala Cys Leu Cys Asp Pro
515 520 525

Gly Leu Tyr Gly Arg Phe Cys His Leu Thr Cys Pro Pro Trp Ala
530 535 540
Phe Gly Pro Gly Cys Ser Glu Glu Cys Gln Cys Val Gln Pro His
545 550 555
Thr Gln Ser Cys Asp Lys Arg Asp Gly Ser Cys Ser Cys Lys Ala
560 565 570
Gly Phe Arg Gly Glu Arg Cys Gln Ala Glu Cys Glu Leu Gly Tyr
575 580 585
Phe Gly Pro Gly Cys Trp Gln Ala Cys Thr Cys Pro Val Gly Val
590 595 600
Ala Cys Asp Ser Val Ser Gly Glu Cys Gly Lys Arg Cys Pro Ala
605 610 615
Gly Phe Gln Gly Glu Asp Cys Gly Gln Glu Cys Pro Val Gly Thr
620 625 630
Phe Gly Val Asn Cys Ser Ser Ser Cys Ser Cys Gly Gly Ala Pro
635 640 645
Cys His Gly Val Thr Gly Gln Cys Arg Cys Pro Pro Gly Arg Thr
650 655 660
Gly Glu Asp Cys Glu Ala Asp Cys Pro Glu Gly Arg Trp Gly Leu
665 670 675
Gly Cys Gln Glu Ile Cys Pro Ala Cys Gln His Ala Ala Arg Cys
680 685 690
Asp Pro Glu Thr Gly Ala Cys Leu Cys Leu Pro Gly Phe Val Gly
695 700 705
Ser Arg Cys Gln Asp Val Cys Pro Ala Gly Trp Tyr Gly Pro Ser
710 715 720
Cys Gln Thr Arg Cys Ser Cys Ala Asn Asp Gly His Cys His Pro
725 730 735
Ala Thr Gly His Cys Ser Cys Ala Pro Gly Trp Thr Gly Phe Ser
740 745 750
Cys Gln Arg Ala Cys Asp Thr Gly His Trp Gly Pro Asp Cys Ser
755 760 765
His Pro Cys Asn Cys Ser Ala Gly His Gly Ser Cys Asp Ala Ile
770 775 780
Ser Gly Leu Cys Leu Cys Glu Ala Gly Tyr Val Gly Pro Arg Cys
785 790 795
Glu Gln Gln Cys Pro Gln Gly His Phe Gly Pro Gly Cys Glu Gln
800 805 810
Leu Cys Gln Cys Gln His Gly Ala Ala Cys Asp His Val Ser Gly
815 820 825
Ala Cys Thr Cys Pro Ala Gly Trp Arg Gly Thr Phe Cys Glu His
830 835 840
Ala Cys Pro Ala Gly Phe Phe Gly Leu Asp Cys Arg Ser Ala Cys
845 850 855
Asn Cys Thr Ala Gly Ala Ala Cys Asp Ala Val Asn Gly Ser Cys
860 865 870
Leu Cys Pro Ala Gly Arg Arg Gly Pro Arg Cys Ala Glu Thr Cys
875 880 885
Pro Ala His Thr Tyr Gly His Asn Cys Ser Gln Ala Cys Ala Cys
890 895 900
Phe Asn Gly Ala Ser Cys Asp Pro Val His Gly Gln Cys His Cys
905 910 915
Ala Pro Gly Trp Met Gly Pro Ser Cys Leu Gln Glu Cys Leu Pro
920 925 930
Arg Asp Val Arg Ala Gly Cys Arg His Ser Gly Gly Cys Leu Asn
935 940 945

Gly Gly Leu Cys Asp Pro His Thr Gly Arg Cys Leu Cys Pro Ala
950 955 960
Gly Trp Thr Gly Asp Lys Cys Gln Ser Pro Cys Leu Arg Gly Trp
965 970 975
Phe Gly Glu Ala Cys Ala Gln Arg Cys Ser Cys Pro Pro Gly Ala
980 985 990
Ala Cys His His Val Thr Gly Ala Cys Arg Cys Pro Pro Gly Phe
995 1000 1005
Thr Gly Ser Gly Cys Glu Gln Gly Cys Pro Pro Gly Arg Tyr Gly
1010 1015 1020
Pro Gly Cys Glu Gln Leu Cys Gly Cys Leu Asn Gly Gly Ser Cys
1025 1030 1035
Asp Ala Ala Thr Gly Ala Cys Arg Cys Pro Thr Gly Phe Leu Gly
1040 1045 1050
Thr Asp Cys Asn Leu Ser Glu Trp Leu Val Ala Ala Val Leu Ser
1055 1060 1065
Gly Ala Ser Cys Val Ser Arg Pro Val Arg Arg Ala Ala Ser Ala
1070 1075 1080
Pro Thr Ala Pro Thr Cys Val Gly Val Gly Arg Gly Arg Pro Ala
1085 1090 1095
Thr Leu

<210> 46
<211> 1350
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503773CD1

<400> 46
Met Gly Ala Ser Arg Asp Arg Gly Leu Ala Ala Leu Trp Cys Leu
1 5 10 15
Gly Leu Leu Gly Gly Leu Ala Arg Val Ala Gly Thr His Tyr Arg
20 25 30
Tyr Leu Trp Arg Gly Cys Tyr Pro Cys His Leu Gly Gln Ala Gly
35 40 45
Tyr Pro Val Ser Ala Gly Asp Gln Arg Pro Asp Val Asp Glu Cys
50 55 60
Arg Thr His Asn Gly Gly Cys Gln His Arg Cys Val Asn Thr Pro
65 70 75
Gly Ser Tyr Leu Cys Glu Cys Lys Pro Gly Phe Arg Leu His Thr
80 85 90
Asp Ser Arg Thr Cys Leu Ala Ile Asn Ser Cys Ala Leu Gly Asn
95 100 105
Gly Gly Cys Gln His His Cys Val Gln Leu Thr Ile Thr Arg His
110 115 120
Arg Cys Gln Cys Arg Pro Gly Phe Gln Leu Gln Glu Asp Gly Arg
125 130 135
His Cys Val Arg Arg Ser Pro Cys Ala Asn Arg Asn Gly Ser Cys
140 145 150
Met His Arg Cys Gln Val Val Arg Gly Leu Ala Arg Cys Glu Cys
155 160 165
His Val Gly Tyr Gln Leu Ala Ala Asp Gly Lys Ala Cys Glu Asp

```
                170                 175                 180
Val Asp Glu Cys Ala Ala Gly Leu Ala Gln Cys Ala His Gly Cys
                185                 190                 195
Leu Asn Thr Gln Gly Ser Phe Lys Cys Val Cys His Ala Gly Tyr
                200                 205                 210
Glu Leu Gly Ala Asp Gly Arg Gln Cys Tyr Arg Ile Glu Met Glu
                215                 220                 225
Ile Val Asn Ser Cys Glu Ala Asn Asn Gly Gly Cys Ser His Gly
                230                 235                 240
Cys Ser His Thr Ser Ala Gly Pro Leu Cys Thr Cys Pro Arg Gly
                245                 250                 255
Tyr Glu Leu Asp Thr Asp Gln Arg Thr Cys Ile Asp Val Asp Asp
                260                 265                 270
Cys Ala Asp Ser Pro Cys Cys Gln Gln Val Cys Thr Asn Asn Pro
                275                 280                 285
Gly Gly Tyr Glu Cys Gly Cys Tyr Ala Gly Tyr Arg Leu Ser Ala
                290                 295                 300
Asp Gly Cys Gly Cys Glu Asp Val Asp Glu Cys Ala Ser Ser Arg
                305                 310                 315
Gly Gly Cys Glu His His Cys Thr Asn Leu Ala Gly Ser Phe Gln
                320                 325                 330
Cys Ser Cys Glu Ala Gly Tyr Arg Leu His Glu Asp Arg Arg Gly
                335                 340                 345
Cys Ser Pro Leu Glu Glu Pro Met Val Asp Leu Asp Gly Glu Leu
                350                 355                 360
Pro Phe Val Arg Pro Leu Pro His Ile Ala Val Leu Gln Asp Glu
                365                 370                 375
Leu Pro Gln Leu Phe Gln Asp Asp Asp Val Gly Ala Asp Glu Glu
                380                 385                 390
Glu Ala Glu Leu Arg Gly Glu His Thr Leu Thr Glu Lys Phe Val
                395                 400                 405
Cys Leu Asp Asp Ser Phe Gly His Asp Cys Ser Leu Thr Cys Asp
                410                 415                 420
Asp Cys Arg Asn Gly Gly Thr Cys Leu Leu Gly Leu Asp Gly Cys
                425                 430                 435
Asp Cys Pro Glu Gly Trp Thr Gly Leu Ile Cys Asn Glu Thr Cys
                440                 445                 450
Pro Pro Asp Thr Phe Gly Lys Asn Cys Ser Phe Ser Cys Ser Cys
                455                 460                 465
Gln Asn Gly Gly Thr Cys Asp Ser Val Thr Gly Ala Cys Arg Cys
                470                 475                 480
Pro Pro Gly Val Ser Gly Thr Asn Cys Glu Asp Gly Cys Pro Lys
                485                 490                 495
Gly Tyr Tyr Gly Lys His Cys Arg Lys Lys Cys Asn Cys Ala Asn
                500                 505                 510
Arg Gly Arg Cys His Arg Leu Tyr Gly Ala Cys Leu Cys Asp Pro
                515                 520                 525
Gly Leu Tyr Gly Arg Phe Cys His Leu Thr Cys Pro Pro Trp Ala
                530                 535                 540
Phe Gly Pro Gly Cys Ser Glu Glu Cys Gln Cys Val Gln Pro His
                545                 550                 555
Thr Gln Ser Cys Asp Lys Arg Asp Gly Ser Cys Ser Cys Lys Ala
                560                 565                 570
Gly Phe Arg Gly Glu Arg Cys Gln Ala Glu Cys Glu Leu Gly Tyr
                575                 580                 585
Phe Gly Pro Gly Cys Trp Gln Ala Cys Thr Cys Pro Val Gly Val
```

590 595 600
Ala Cys Asp Ser Val Ser Gly Glu Cys Gly Lys Arg Cys Pro Ala
605 610 615
Gly Phe Gln Gly Glu Asp Cys Gly Gln Glu Cys Pro Val Gly Thr
620 625 630
Phe Gly Val Asn Cys Ser Ser Ser Cys Ser Cys Gly Gly Ala Pro
635 640 645
Cys His Gly Val Thr Gly Gln Cys Arg Cys Pro Pro Gly Arg Thr
650 655 660
Gly Glu Asp Cys Glu Ala Asp Cys Pro Glu Gly Arg Trp Gly Leu
665 670 675
Gly Cys Gln Glu Ile Cys Pro Ala Cys Gln His Ala Ala Arg Cys
680 685 690
Asp Pro Glu Thr Gly Ala Cys Leu Cys Leu Pro Gly Phe Val Gly
695 700 705
Ser Arg Cys Gln Asp Val Cys Pro Ala Gly Trp Tyr Gly Pro Ser
710 715 720
Cys Gln Thr Arg Cys Ser Cys Ala Asn Asp Gly His Cys His Pro
725 730 735
Ala Thr Gly His Cys Ser Cys Ala Pro Gly Trp Thr Gly Phe Ser
740 745 750
Cys Gln Arg Ala Cys Asp Thr Gly His Trp Gly Pro Asp Cys Ser
755 760 765
His Pro Cys Asn Cys Ser Ala Gly His Gly Ser Cys Asp Ala Ile
770 775 780
Ser Gly Leu Cys Leu Cys Glu Ala Gly Tyr Val Gly Pro Arg Cys
785 790 795
Glu Gln Gln Cys Pro Gln Gly His Phe Gly Pro Gly Cys Glu Gln
800 805 810
Leu Cys Gln Cys Gln His Gly Ala Ala Cys Asp His Val Ser Gly
815 820 825
Ala Cys Thr Cys Pro Ala Gly Trp Arg Gly Thr Phe Cys Glu His
830 835 840
Ala Cys Pro Ala Gly Phe Phe Gly Leu Asp Cys Arg Ser Ala Cys
845 850 855
Asn Cys Thr Ala Gly Ala Ala Cys Asp Ala Val Asn Gly Ser Cys
860 865 870
Leu Cys Pro Ala Gly Arg Arg Gly Pro Arg Cys Ala Glu Thr Cys
875 880 885
Pro Ala His Thr Tyr Gly His Asn Cys Ser Gln Ala Cys Ala Cys
890 895 900
Phe Asn Gly Ala Ser Cys Asp Pro Val His Gly Gln Cys His Cys
905 910 915
Ala Pro Gly Trp Met Gly Pro Ser Cys Leu Gln Glu Cys Leu Pro
920 925 930
Arg Asp Val Arg Ala Gly Cys Arg His Ser Gly Gly Cys Leu Asn
935 940 945
Gly Gly Leu Cys Asp Pro His Thr Gly Arg Cys Leu Cys Pro Ala
950 955 960
Gly Trp Thr Gly Asp Lys Cys Gln Ser Pro Cys Leu Arg Gly Trp
965 970 975
Phe Gly Glu Ala Cys Ala Gln Arg Cys Ser Cys Pro Pro Gly Ala
980 985 990
Ala Cys His His Val Thr Gly Ala Cys Arg Cys Pro Pro Gly Phe
995 1000 1005
Thr Gly Ser Gly Cys Glu Gln Ala Cys Pro Pro Gly Ser Phe Gly

1010 1015 1020
Glu Asp Cys Ala Gln Met Cys Gln Cys Pro Gly Glu Asn Pro Ala
1025 1030 1035
Cys His Pro Ala Thr Gly Thr Cys Ser Cys Ala Ala Gly Tyr His
1040 1045 1050
Gly Pro Ser Cys Gln Gln Arg Cys Pro Pro Gly Arg Tyr Gly Pro
1055 1060 1065
Gly Cys Glu Gln Leu Cys Gly Cys Leu Asn Gly Gly Ser Cys Asp
1070 1075 1080
Ala Ala Thr Gly Ala Cys Arg Cys Pro Thr Gly Phe Leu Gly Thr
1085 1090 1095
Asp Cys Asn Leu Thr Cys Pro Gln Gly Arg Phe Gly Pro Asn Cys
1100 1105 1110
Thr His Val Cys Gly Cys Gly Gln Gly Ala Ala Cys Asp Pro Val
1115 1120 1125
Thr Gly Thr Cys Leu Cys Pro Pro Gly Arg Ala Gly Val Arg Cys
1130 1135 1140
Glu Arg Gly Cys Pro Gln Asn Arg Phe Gly Val Gly Cys Glu His
1145 1150 1155
Thr Cys Ser Cys Arg Asn Gly Gly Leu Cys His Ala Ser Asn Gly
1160 1165 1170
Ser Cys Ser Cys Gly Leu Gly Trp Thr Gly Arg His Cys Glu Leu
1175 1180 1185
Ala Cys Pro Pro Gly Arg Tyr Gly Ala Ala Cys His Leu Glu Cys
1190 1195 1200
Ser Cys His Asn Asn Ser Thr Cys Glu Pro Ala Thr Gly Thr Cys
1205 1210 1215
Arg Cys Gly Pro Gly Phe Tyr Gly Gln Ala Cys Glu His Pro Cys
1220 1225 1230
Pro Pro Gly Phe His Gly Ala Gly Cys Gln Gly Leu Cys Trp Cys
1235 1240 1245
Gln His Gly Ala Pro Cys Asp Pro Ile Ser Gly Arg Cys Leu Cys
1250 1255 1260
Pro Ala Gly Phe His Gly His Phe Cys Glu Arg Gly Cys Glu Pro
1265 1270 1275
Gly Ser Phe Gly Glu Gly Cys His Gln Arg Cys Asp Cys Asp Gly
1280 1285 1290
Gly Ala Pro Cys Asp Pro Val Thr Gly Leu Cys Leu Cys Pro Pro
1295 1300 1305
Gly Arg Ser Gly Ala Thr Cys Asn Leu Gly Gly Pro Leu Arg Leu
1310 1315 1320
Pro Glu Asn Pro Ser Leu Ala Gln Gly Ser Ala Gly Thr Leu Pro
1325 1330 1335
Ala Ser Ser Arg Pro Thr Ser Arg Ser Gly Gly Pro Ala Arg His
1340 1345 1350

<210> 47
<211> 71
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504698CD1

<400> 47
Met Glu Val Pro Pro Pro Ala Pro Arg Ser Phe Leu Cys Arg Ala
1 5 10 15
Leu Cys Leu Phe Pro Arg Val Phe Ala Ala Glu Ala Val Thr Ala
20 25 30
Asp Ser Glu Val Leu Glu Glu Arg Gln Lys Arg Leu Pro Tyr Val
35 40 45
Pro Glu Pro Tyr Tyr Pro Glu Ser Gly Trp Asp Arg Leu Arg Glu
50 55 60
Leu Phe Gly Lys Asp Cys His Gly Lys Ser Phe
65 70

<210> 48
<211> 220
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510361CD1

<400> 48
Met Leu Gln Leu Trp Lys Leu Val Leu Leu Cys Gly Val Leu Thr
1 5 10 15
Gly Thr Ser Glu Ser Leu Leu Asp Asn Leu Gly Asn Asp Leu Ser
20 25 30
Asn Val Val Asp Lys Leu Glu Pro Val Leu His Glu Gly Leu Glu
35 40 45
Thr Val Asp Asn Thr Leu Lys Gly Ile Leu Glu Lys Leu Lys Val
50 55 60
Asp Leu Gly Val Leu Gln Lys Ser Ser Ala Trp Gln Leu Ala Lys
65 70 75
Gln Lys Ala Gln Glu Ala Glu Lys Leu Leu Asn Asn Val Ile Ser
80 85 90
Lys Leu Leu Pro Thr Asn Thr Asp Ile Phe Gly Leu Lys Ile Ser
95 100 105
Asn Ser Leu Ile Leu Asp Val Lys Ala Glu Pro Ile Asp Asp Gly
110 115 120
Lys Gly Leu Asn Leu Ser Phe Pro Val Thr Ala Asn Val Thr Val
125 130 135
Ala Gly Pro Ile Ile Gly Gln Ile Ile Asn Leu Lys Ala Ser Leu
140 145 150
Asp Leu Leu Thr Ala Val Thr Ile Glu Thr Asp Pro Gln Thr His
155 160 165
Gln Pro Val Ala Val Leu Gly Glu Cys Ala Ser Asp Pro Thr Ser
170 175 180
Ile Ser Leu Ser Leu Leu Asp Asn Gln Lys Cys Ile Glu Ala Gly
185 190 195
His Asp Gly Ser Thr Pro Val Ile Pro Ala Leu Trp Glu Ala Glu
200 205 210
Thr Gln Pro Asn His Gln Gln Val Arg Glu
215 220

<210> 49
<211> 39
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507013CD1

<400> 49
Met Arg Gly His Pro Ser Leu Leu Leu Leu Tyr Met Ala Leu Thr
1 5 10 15
Thr Cys Leu Asp Thr Ser Pro Ser Glu Glu Thr Asp Gln Glu Val
20 25 30
Phe Leu Gly Arg Gly Gln Gly Val Phe
35

<210> 50
<211> 451
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510507CD1

<400> 50
Met Ala Ser Ala Thr Glu Asp Pro Val Leu Glu Arg Tyr Phe Lys
1 5 10 15
Gly His Lys Ala Ala Ile Thr Ser Leu Asp Leu Ser Pro Asn Gly
20 25 30
Lys Gln Leu Ala Thr Ala Ser Trp Asp Thr Phe Leu Met Leu Trp
35 40 45
Asn Phe Lys Pro His Ala Arg Ala Tyr Arg Tyr Val Gly His Lys
50 55 60
Asp Val Val Thr Ser Val Gln Phe Ser Pro His Gly Asn Leu Leu
65 70 75
Ala Ser Ala Ser Arg Asp Arg Thr Val Arg Leu Trp Ile Pro Asp
80 85 90
Lys Arg Gly Lys Phe Ser Glu Phe Lys Ala His Thr Ala Pro Val
95 100 105
Arg Ser Val Asp Phe Ser Ala Asp Gly Gln Phe Leu Ala Thr Ala
110 115 120
Ser Glu Asp Lys Ser Ile Lys Val Trp Ser Met Tyr Arg Gln Arg
125 130 135
Phe Leu Tyr Ser Leu Tyr Arg His Thr His Trp Val Arg Cys Ala
140 145 150
Lys Phe Ser Pro Asp Gly Arg Leu Ile Val Ser Cys Ser Glu Asp
155 160 165
Lys Thr Ile Lys Ile Trp Asp Thr Thr Asn Lys Gln Cys Val Asn
170 175 180
Asn Phe Ser Asp Phe Val Gly Phe Ala Asn Phe Val Asp Phe Asn
185 190 195
Pro Ser Gly Thr Cys Ile Ala Ser Ala Gly Ser Asp Gln Thr Val
200 205 210
Lys Val Trp Asp Val Arg Val Asn Lys Leu Leu Gln His Tyr Gln
215 220 225
Val His Ser Gly Gly Val Asn Cys Ile Ser Phe His Pro Ser Gly
230 235 240

```
Asn Tyr Leu Ile Thr Ala Ser Ser Asp Gly Thr Leu Lys Ile Leu
                245                 250                 255
Asp Leu Leu Glu Gly Arg Leu Ile Tyr Thr Leu Gln Gly His Thr
                260                 265                 270
Gly Pro Val Phe Thr Val Ser Phe Ser Lys Gly Gly Glu Leu Phe
                275                 280                 285
Ala Ser Gly Gly Ala Asp Thr Gln Val Leu Leu Trp Arg Thr Asn
                290                 295                 300
Phe Asp Glu Leu His Cys Lys Gly Leu Thr Lys Arg Asn Leu Lys
                305                 310                 315
Arg Leu His Phe Asp Ser Pro Pro His Leu Leu Asp Ile Tyr Pro
                320                 325                 330
Arg Thr Pro His Pro His Glu Glu Lys Val Glu Thr Val Glu Thr
                335                 340                 345
Thr Glu Thr Ser Gly Arg Thr Leu Pro Asp Lys Gly Glu Glu Ala
                350                 355                 360
Cys Gly Tyr Phe Leu Asn Pro Ser Leu Met Ser Pro Glu Cys Leu
                365                 370                 375
Pro Thr Thr Thr Lys Lys Lys Thr Glu Asp Met Ser Asp Leu Pro
                380                 385                 390
Cys Glu Ser Gln Arg Ser Ile Pro Leu Ala Val Thr Asp Ala Leu
                395                 400                 405
Glu His Ile Met Glu Gln Leu Asn Val Leu Thr Gln Thr Val Ser
                410                 415                 420
Ile Leu Glu Gln Arg Leu Thr Leu Thr Glu Asp Lys Leu Lys Asp
                425                 430                 435
Cys Leu Glu Asn Gln Gln Lys Leu Phe Ser Ala Val Gln Gln Lys
                440                 445                 450
Ser


<210> 51
<211> 224
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 90106370CD1

<400> 51
Met Pro Leu Thr Pro Glu Pro Pro Ser Gly Arg Val Glu Gly Pro
  1               5                  10                  15
Pro Ala Trp Glu Ala Ala Pro Trp Pro Ser Leu Pro Cys Gly Pro
                 20                  25                  30
Cys Ile Pro Ile Met Leu Val Leu Ala Thr Leu Ala Ala Leu Phe
                 35                  40                  45
Ile Leu Thr Thr Ala Val Leu Ala Glu Arg Leu Phe Arg Arg Ala
                 50                  55                  60
Leu Arg Pro Asp Pro Ser His Arg Ala Pro Thr Leu Val Trp Arg
                 65                  70                  75
Pro Gly Gly Glu Leu Trp Ile Glu Pro Met Gly Thr Ala Arg Glu
                 80                  85                  90
Arg Ser Glu Asp Trp Tyr Gly Ser Ala Val Pro Leu Leu Thr Asp
                 95                 100                 105
Arg Ala Pro Glu Pro Pro Thr Gln Val Gly Thr Leu Glu Ala Arg
```

```
                110                 115                 120
Ala Thr Ala Pro Pro Ala Pro Ser Ala Pro Asn Ser Ala Pro Ser
                125                 130                 135
Asn Leu Gly Pro Gln Thr Val Leu Glu Val Pro Ala Arg Ser Thr
                140                 145                 150
Phe Trp Gly Pro Gln Pro Trp Glu Gly Arg Pro Pro Ala Thr Gly
                155                 160                 165
Leu Val Ser Trp Ala Glu Pro Glu Gln Arg Pro Glu Ala Ser Val
                170                 175                 180
Gln Phe Gly Ser Pro Gln Ala Arg Arg Gln Arg Pro Gly Ser Pro
                185                 190                 195
Asp Pro Glu Trp Gly Leu Gln Pro Arg Val Thr Leu Glu Gln Ile
                200                 205                 210
Ser Ala Phe Trp Lys Arg Glu Gly Arg Thr Ser Val Gly Phe
                215                 220
```

<210> 52
<211> 1480
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506904CB1

<400> 52

```
aatctgcaaa gttcctgtga gcgctgtcat tttgtcactc tggtttttca gattcttccc 60
ctggaggctg gagtttccag gatgtcaaaa ttacctctgc ttgggtgagc tatttcaagc 120
agctgggata cctgtgtcac tcctgctgtc tgccagtgac tgcccaggtg tctgctggtt 180
cctccccagg agtagggagg aaccaggtgg gctggctggg atgggtggat atttaaagac 240
caggccttgg acgctgcagc acttctatct ctgcttgatg cctgctgcca cgtggctggt 300
cctcctcctc ctgctgtggc tgagccttgg ggtgaagaca ggcagctgct cccaacccca 360
gaacctttgc tgtcttggga cggatcacca ctgcaagagg ggaagttgct actgtgatga 420
attctgccat gtggcaccag actgccaccc agaccacagt gtcctctgca accctggtaa 480
ctcacataca ggcccgattc cacctacagc aaagctggat gcgatggctg gcagaggcaa 540
accctttgcc tgcacttcag gccaaagccg ggatgtggcc tagatggttc ctaaggtccc 600
tgacaatcct gagatcttgc atcttgtcta tttcaggtca aagcttctca gatgaccaag 660
atggtgctgc agatggtgct gaggatggag aacccaccaa gccccgctag gagccaccta 720
gactggatgc agagcatggt gagctccctg caggttctct gagaaggggt ggatggcagc 780
ctgctccttg cctttgtgcc ctccaggccc caaagtcagg gaaccaaaag aagaaagggg 840
ccgtagctag ggcagagctc cactgcaatg attgttttag ggtaggagc caggattgcc 900
gtctgtggac actgaaattt gaatctcata tacttttgtg acaaaacatt cttcctcttt 960
tgttcttctc ctaccatcta aaaatgtaga aaacattctt agcctatgag ttgcacaaaa 1020
acaggcagtg gccagatttg gcccatagac catagtttgc tgacttctgc cctaaatcat 1080
cctccatttc tttccttctg tgtccttgtt actgacaaag ccactttccc taaaatgggg 1140
tctttccctg tttggtgcca tgaagccaat atgcaaaacc gaaagtgagc ctcaagcagt 1200
gcaggcttta tttggtggcc atggaattga gaagtgagag cttggctcac aaatcaactt 1260
ttctgctcgt gagacccagg aagtcacaga tacagggcat ctttagtgaa ggggctgagc 1320
attaaaagca aggggaggag tggccaggtg caatggctca ctcccataaa cccagaactt 1380
tgggaggcca aaatgagagg attgctgaga ccaggagttc gagaccatcc tggtcaacat 1440
agtgatacac ccccatctct acaaaaataa aaaaaaaaaa                       1480
```

<210> 53
<211> 1168
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506909CB1

<400> 53
aatctgcaaa gttcctgtga gcgctgtcat tttgtcactc tggtttttca gattcttccc 60
ctggaggctg gagtttccag gatgtcaaaa ttacctctgc ttgggtgagc tatttcaagc 120
agctgggata cctgtgtcac tcctgctgtc tgccagtgac tgcccaggtg tctgctggtt 180
cctccccagg agtagggagg aaccaggtgg gctggctggg atgggtggat atttaaagac 240
caggccttgg acgctgcagc acttctatct ctgcttgatg cctgctgcca cgtggctggt 300
cctcctcctc ctgctgtggc tgagccttgg ggtgaagaca gcttctcaga tgaccaagat 360
ggtgctgcag atggtgctga ggatggagaa cccaccaagc cccgctagga gccacctaga 420
ctggatgcag agcatggtga gctccctgca ggttctctga gaaggggtgg atggcagcct 480
gctccttgcc tttgtgccct ccaggcccca aagtcaggga accaaaagaa gaaaggggcc 540
gtagctaggg cagagctcca ctgcaatgat tgttttaggg gtaggagcca ggattgccgt 600
ctgtggacac tgaaatttga atctcatata cttttgtgac aaaacattct tcctcttttg 660
ttcttctcct accatctaaa aatgtagaaa acattcttag cctatgagtt gcacaaaaac 720
aggcagtggc cagatttggc ccatagacca tagtttgctg acttctgccc taaatcatcc 780
tccatttctt tccttctgtg tccttgttac tgacaaagcc actttcccta aaatggggtc 840
tttccctgtt tggtgccatg aagccaatat gcaaaaccga aagtgagcct caagcagtgc 900
aggctttatt tggtggccat ggaattgaga agtgagagct tggctcacaa atcaactttt 960
ctgctcgtga gacccaggaa gtcacagata cagggcatct ttagtgaagg ggctgagcat 1020
taaaagcaag gggaggagtg gccaggtgca atggctcact cccataaacc cagaactttg 1080
ggaggccaaa atgagaggat tgctgagacc aggagttcga gaccatcctg gtcaacatag 1140
tgatacaccc ccatctctac aaaaataa 1168

<210> 54
<211> 2229
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507096CB1

<400> 54
agactctgcg cacgcgcagt tccccctccc gcccctgctt gccggtgatg gcgcataacg 60
catgcgcggg gagggcggag ctgggcgttg ccgtggctac tgggaacgca tttcacgggg 120
gcggggcgtg gttccggggc ggggcgcggc cgccggaagt gcgtggccgc ccggggccat 180
ggcgacactc agcttcgtct tcctgctgct gggggcagtg tcctggcctc cggcttctgc 240
ctccggccag gagttctggc ccggacaatc ggcggccgat attctgtcgg gggcggcttc 300
ccgcagacgg tatcttctgt atgacgtcaa ccccccggaa ggcttcaacc tgcgcaggga 360
tgtctatatc cgaatcgcct ctctcctgaa gactctgctg aagacggagg agtgggtgct 420
tgtcctgcct ccatggggcc gcctctatca ctggcagagt cctgacatcc accaggtccg 480
gattccctgg tctgagtttt ttgatcttcc aagtctcaat aaaaacatcc ccgtcatcga 540
gtatgagcag ttcatcgcag ggctcagcct ccatcgtggc gcccctgctg ctgagaaaca 600
catcagcccg gtccgtgatg ttagacagag ccgagaacct acttcacgac cactatggag 660
ggaaagaata ctgggatacc cgtcgcagca tggtgtttgc caggcacctg cgggaggtgg 720
gagacgagtt caggagcaga catctcaact ccacggacga cgcagacagg atccccttcc 780
aggaggactg gatgaagatg aaggtcaagc tgggctccgc gctagggggc ccctacctgg 840
gagtccacct gagaagaaaa gatttcatct ggggtcacag acaggatgta cccagtctgg 900
aaggggccgt gaggaagatc cgcagcctca tgaagaccca ccggctggac aaggtgtttg 960
tggccacaga tgccgtcaga aaggaatatg aagagctaaa aaagctgtta cccgagatgg 1020
tgaggtttga acccacgtgg gaggagctgg agctctacaa ggacggaggc gttgcgatta 1080
ttgaccagtg gatctgcgca cacgccaggt tttttattgg cacctcagtc tcaacatttt 1140
cttttcggat tcatgaggaa agagaaatcc tggggttgga ccccaagacg acgtacaaca 1200

ggttctgcgg agaccaagag aaggcgtgtg agcaacccac ccactggaag atcacctact 1260
gaggaggatc ctccagggcc gctccccgga cccgacaggc gcgggtggat gcaggttctg 1320
tcgccgtgga gtcaccgtct actgccagcc gggagctggg cggacaggac cgtccctcgc 1380
agggtcccag gcccagaaga ggccccacgc ctctagagct gggctccgtc ctcggcgttg 1440
ccagccgcca tggctgatga agaggctccg ctgctctcgg gggtggcggt tgttttcagg 1500
cagcgtctgt gaacccacag ctcggttgcc agcagtgccc gcgtggtgac ccagaagcag 1560
gagtgtttgt caggctcccg ctctggcctt tccagccacc tttcatgtct tcatatttta 1620
agtgcattga ggatagatgc aggcgggtga gctgccctcc gtcaggtgga cccgggctga 1680
catttccctg ggagctggtg caaggagaag cgtcatttta aatgtctgca gagcgaccag 1740
gggcctcatg aatctctccg ttgccctccg cgcagcagga ggctgcctgt gtgtttcctc 1800
ctgggatgcg tgcaaggcag acctggtgct gcaaaggaaa gggcctgagg cctcagggag 1860
ccccgtggag ggatgacagt tcaggcccta ctgctggcac gtcagagcac tgggaagttt 1920
ttcagtgacg tctctggggc actcagtgga ttgtctgtag gaaacttgca gctctgctcc 1980
tcacaccagg cccggctggc cacccaccct cgcccccact ggccacccct ccctcgcccc 2040
gactgccccg ccccaccctc accccgactg ccccgccctc gcccggctgg ccgtccctgc 2100
cctcgccccg gctggcaggt gcacatgggg cctccaggtc tgccattcgc tattgagaac 2160
tagaaatgag gaaggacagt tacgctaact ccaaaaggct gtctaggatg agctgcttta 2220
tcagggagc 2229

<210> 55
<211> 2374
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507098CB1

<400> 55
agactctgcg cacgcgcagt tccccctccc gcccctgctt gccggtgatg gcgcataacg 60
catgcgcggg gagggcggag ctgggcgttg ccgtggctac tgggaacgca tttcacgggg 120
gcggggcgtg gttccggggc ggggcgcggc cgccggaagt gcgtggccgc ccggggccat 180
ggcgacactc agcttcgtct tcctgctgct gggggcagtg tcctggcctc cggcttctgc 240
ctccggccag gagttctggc ccggacaatc ggcggccgat attctgtcgg gggcggcttc 300
ccgcagacgg tatcttctgt atgacgtcaa ccccccggaa ggcttcaacc tgcgcaggga 360
tgtctatatc cgaatcgcct ctctcctgaa gactctgctg aagacggagg agtgggtgct 420
tgtcctgcct ccatggggcc gcctctatca ctggcagagt cctgacatcc accaggtccg 480
gattccctgg tctgagtttt ttgatcttcc aagtctcaat aaaaacatcc ccgtcatcga 540
gtatgagcag ttcatcgcag aatctggtgg gtcctttatt gaccaggttt acgtcctgca 600
aagttacgca gaggggtgga aagaagggac ctgggaagag aaggtggacg agcggccgtg 660
tattgatcag ctcctgtact cccaggacaa gcacgagtac tacagggctc agcctccatc 720
gtggcgcccc tgctgctgag aaacacatca gcccggtccg tgatgttaga cagagccgag 780
aacctacttc acgaccacta tggagggaaa gaatactggg atacccgtcg cagcatggtg 840
tttgccaggc acctgcggga ggtgggagac gagttcagga gcagacatct caactccacg 900
gacgacgcag acaggatccc cttccaggag gactggatga agatgaaggt caagctgggc 960
tccgcgctag gggcccccta cctgggagtc cacctgagaa gaaaagattt catctggggt 1020
cacagacagg atgtacccag tctggaaggg gccgtgagga agatccgcag cctcatgaag 1080
acccaccggc tggacaaggt gtttgtggcc acagatgccg tcagaaagga atatgaagag 1140
ctaaaaaagc tgttacccga gatggtgagg tttgaaccca cgtgggagga gctggagctc 1200
tacaaggacg gaggcgttgc gattattgac cagtggatct gcgcacacgc caggtttttt 1260
attggcacct cagtctcaac attttctttt cggattcatg aggaaagaga aatcctgggg 1320
ttggacccca agacgacgta caacaggttc tgcggagacc aagagaaggc gtgtgagcaa 1380
cccacccact ggaagatcac ctactgagga ggatcctcca gggccgctcc ccggacccga 1440
caggcgcggg tggatgcagg ttctgtcgcc gtggagtcac cgtctactgc cagccgggag 1500
ctgggcggac aggaccgtcc ctcgcagggt cccaggccca gaagaggccc cacgcctcta 1560
gagctgggct ccgtcctcgg cgttgccagc cgccatggct gatgaagagg ctccgctgct 1620

ctcgggggtg gcggttgttt tcaggcagcg tctgtgaacc cacagctcgg ttgccagcag 1680
tgcccgcgtg gtgacccaga agcaggagtg tttgtcaggc tcccgctctg gcctttccag 1740
ccacctttca tgtcttcata ttttaagtgc attgaggata gatgcaggcg ggtgagctgc 1800
cctccgtcag gtggacccgg gctgacattt ccctgggagc tggtgcaagg agaagcgtca 1860
ttttaaatgt ctgcagagcg accaggggcc tcatgaatct ctccgttgcc ctccgcgcag 1920
caggaggctg cctgtgtgtt tcctcctggg atgcgtgcaa ggcagacctg gtgctgcaaa 1980
ggaaagggcc tgaggcctca gggagccccg tggagggatg acagttcagg ccctactgct 2040
ggcacgtcag agcactggga agtttttcag tgacgtctct ggggcactca gtggattgtc 2100
tgtaggaaac ttgcagctct gctcctcaca ccaggcccgg ctggccaccc accctcgccc 2160
ccactggcca cccctccctc gccccgactg ccccgcccca ccctcacccc gactgccccg 2220
ccctcgcccg gctggccgtc cctgccctcg ccccggctgg caggtgcaca tggggcctcc 2280
aggtctgcca ttcgctattg agaactagaa atgaggaagg acagttacgc taactccaaa 2340
aggctgtcta ggatgagctg ctttatcagg gagc 2374

<210> 56
<211> 2151
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507099CB1

<400> 56
agactctgcg cacgcgcagt tccccctccc gcccctgctt gccggtgatg gcgcataacg 60
catgcgcggg gagggcggag ctgggcgttg ccgtggctac tgggaacgca tttcacgggg 120
gcggggcgtg gttccggggc ggggcgcggc cgccggaagt gcgtggccgc ccggggccat 180
ggcgacactc agcttcgtct tcctgctgct gggggcagtg tcctggcctc cggcttctgc 240
ctccggccag gagttctggc ccggacaatc ggcggccgat attctgtcgg gggcggcttc 300
ccgcagacgg tatcttctgt atgacgtcaa ccccccggaa ggcttcaacc tgcgcaggga 360
tgtctatatc cgaatcgcct ctctcctgaa gactctgctg aagacggagg agtgggtgct 420
tgtcctgcct ccatggggcc gcctctatca ctggcagagt cctgacatcc accaggtccg 480
gattccctgg tctgagtttt ttgatcttcc aagtctcaat aaaaacatcc ccgtcatcga 540
gtatgagcag ttcatcgcag aatctggtgg gccctttatt gaccaggttt acgtcctgca 600
aagttacgca gaggggtgga aagaagggac ctgggaagag aaggtggacg agcggccgtg 660
tattgatcaa ctccacggac gacgcagaca ggatcccctt ccaggaggac tggatgaaga 720
tgaaggtcaa gctgggctcc gcgctagggg gcccctacct gggagtccac ctgagaagaa 780
aagatttcat ctggggtcac agacaggatg tacccagtct ggaaggggcc gtgaggaaga 840
tccgcagcct catgaagacc caccggctgg acaaggtgtt tgtggccaca gatgccgtca 900
gaaaggaata tgaagagcta aaaaagctgt tacccgagat ggtgaggttt gaacccacgt 960
gggaggagct ggagctctac aaggacggag gcgttgcgat tattgaccag tggatctgcg 1020
cacacgccag gtttttttatt ggcacctcag tctcaacatt ttcttttcgg attcatgagg 1080
aaagagaaat cctggggttg gaccccaaga cgacgtacaa caggttctgc ggagaccaag 1140
agaaggcgtg tgagcaaccc acccactgga agatcaccta ctgaggagga tcctccaggg 1200
ccgctccccg gacccgacag gcgcgggtgg atgcaggttc tgtcgccgtg gagtcaccgt 1260
ctactgccag ccgggagctg ggcggacagg accgtccctc gcagggtccc aggcccagaa 1320
gaggccccac gcctctagag ctgggctccg tcctcggcgt tgccagccgc catggctgat 1380
gaagaggctc cgctgctctc gggggtggcg gttgttttca ggcagcgtct gtgaacccac 1440
agctcggttg ccagcagtgc ccgcgtggtg acccagaagc aggagtgttt gtcaggctcc 1500
cgctctggcc tttccagcca cctttcatgt cttcatattt taagtgcatt gaggatagat 1560
gcaggcgggt gagctgccct ccgtcaggtg gacccgggct gacatttccc tgggagctgg 1620
tgcaaggaga agcgtcattt taaatgtctg cagagcgacc aggggcctca tgaatctctc 1680
cgttgccctc cgcgcagcag gaggctgcct gtgtgtttcc tcctgggatg cgtgcaaggc 1740
agacctggtg ctgcaaagga aagggcctga ggcctcaggg agccccgtgg agggatgaca 1800
gttcaggccc tactgctggc acgtcagagc actgggaagt tttcagtga cgtctctggg 1860
gcactcagtg gattgtctgt aggaaacttg cagctctgct cctcacacca ggcccggctg 1920

gccacccacc ctcgccccca ctggccaccc ctccctcgcc ccgactgccc cgccccaccc 1980
tcaccccgac tgccccgccc tcgcccggct ggccgtccct gccctcgccc cggctggcag 2040
gtgcacatgg ggcctccagg tctgccattc gctattgaga actagaaatg aggaaggaca 2100
gttacgctaa ctccaaaagg ctgtctagga tgagctgctt tatcagggag c 2151

<210> 57
<211> 3435
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7501399CB1

<400> 57
ggcgagcccg agccgccgcc ggccccgcgg cggagatgag caggtccgcg acgctgctgc 60
tgtgcctgct gggctgccac gtctggaagg cggtgaccaa gacgctgcgg gagcccggcg 120
ccggagccca agaggtgacg ttaaaggtgc acatcagcga cgccagcacc caccagcccg 180
tagcagatgc gctcatcgag atcttcacca accaggcctc catagcctct ggcacctcgg 240
ggactgatgg cgtcgccttt atcaagttcc agtataagct gggcagtcag ttgattgtca 300
ccgcctcgaa gcatgcctac gtgccaaact ctgccccatg gaagccaatc cggttacctg 360
tattttcctc tctgagcctt ggcctgcttc cagaacgctc tgccactcta atggtatatg 420
aagatgtcgt ccaaatagta tcaggattcc aaggtgcccg gccacagcct cgcgttcatt 480
tccagagaag ggctctgagg ttgcctgaga acaccagcta cagtgacctg accgcgtttc 540
tcacggccgc cagctcccct tcggaggtgg acagttttcc ttatttgcga ggattagacg 600
gaaatggaac aggaaacagc accaggcatg acctgacccc agtcacagcc gtcagcgtcc 660
acttgctgag cagtaatgga acgccggtgc tggtggatgg tcccatctat gtcactgtgc 720
ccctggccac gcagagcagc ctgaggcaca atgcctatgt cgcggcgtgg cggtttgacc 780
agaagctggg aacgtggctg aagagcggtc tgggtcttgt gcaccaggaa ggcagccagc 840
tgacgtggac atacattgcc ccccagttgg ggtactgggt ggccgccatg tcccctccca 900
tcccaggagg aagtgcttga aacctcgtca gcaccacaga aaactgcagc tccctgcagg 960
actggagagt tccaaaagag accagtccac gtccatgtca cacattaact tgctgttttc 1020
acgccgagcg tcagaattcc ctggcccgct gtccgtcacc agccacggcc gccccgaggc 1080
ccccggcacg aaggaactga tgagtggagt ccatttggaa atgatgtctc cgggcggcga 1140
aggggacctg cacacccccca tgctcaagct ctcctacagc acctcccagg aatttagctc 1200
ccgggaggag ctcctctctt gcaaggaaga ggataaaagc cagatctcct ttgataacct 1260
cactccaagt gggacgctgg ggaaagacta ccataagtca gtggaggttt ttcccttaaa 1320
ggcaagaaaa tctatggaaa gagaaggcta cgagtcctcg ggcaatgatg actacagggg 1380
tagttacaac accgtgctct cacagccttt atttgaaaag caggacagag aaggtccagc 1440
ctccacggga agcaaactca ccattcagga acatctgtac ccccgcgcctt catcacctga 1500
gaaagaacag ctgctggacc gcagacccac tgaatgtatg atgtcgcgat cagtagatca 1560
cctcgagaga cctacgtcct tcccacggcc cggccagtta atctgctgca gttctgtcga 1620
ccaggtcaat gacagcgttt acaggaaagt actgcctgcc ttggtcatcc cggctcatta 1680
tatgaaactc cccggggacc actcctatgt cagccagccc ctcgtcgtcc cggctgatca 1740
gcagcttgag atagaaagac tacaggctga gctgtccaat ccccatgccg ggatcttccc 1800
acacccgtcc tcacagatcc agccccagcc cctgtcttcc caggccatct ctcagcagca 1860
cctgcaggat gcgggcaccc gggagtggag ccctcagaac gcatccatgt cggagtctct 1920
ctccatccca gcttccctga acgacgcggc tttggctcag atgaacagtg aggtgcagct 1980
cctgactgaa aaggccctga tggagcttgg gggtgggaag ccgcttccgc acccccggc 2040
gtggttcgtc tccttggatg gcaggtccaa cgctcacgtt agacattcat acattgatct 2100
ccaaagagct ggaaggaacg gaagtaatga tgccagtttg gactctggcg tagatatgaa 2160
tgaaccaaaa tcagcccgga agggaagggg agatgctttg tctctgcagc agaactaccc 2220
gcccgtccaa gagcaccagc agaaagagcc tcgagcccca gacagcacgg cctacacgca 2280
gctcgtgtac ctggatgacg tggaacagag tggtagcgaa tgtgggacca cggtctgtac 2340
cccgaggac agtgccctgc gatgcttgtt ggaggggtcg agtcggagaa gtggtggcca 2400
gctgcccagc ctgcaggagg agacgaccag acggactgcg gatgcccccct cggagccagc 2460

agccagcccc caccagagaa gatctgccca cgaggaagag gaagacgatg atgatgatga 2520
ccaaggagaa gacaagaaaa gcccctggca gaaacgggag gagaggcccc tgatggcgtt 2580
taacattaaa tgagctatcg cagacccacc tgactgtgga atataaaatt gccaaatatc 2640
ctttctcatg gaagcgcgta cccgttcgtg gaggaaacgg aacggcagcc cagccgtggg 2700
acggacgtgg acgtttactg cattcctgtt tgccgtgtaa atgttagaaa ggaattaaag 2760
ttattactcg gaataaagga tgactttggc ggatgtcgcc cctgcaagga ggtggctgaa 2820
agtggtgtcc agatgtcctt ccgaggactc ggcgtatccg ccaccaggga cattaagaaa 2880
ccgcacgtga tgtcgctatg ctctaacgat cacctcagtt ctccctcgga ttctgggaac 2940
agatgaaact ttttgcatcg cttgagtcat ttttatcaca ataatcctac tgtgaagctg 3000
tcgttgagaa cttaggttgg cacgtagcgt ctcaaggtat gcgttctctc aaaggaaagc 3060
tatgcatcgc tgcttcgttg tctgatttttg cttagatttt gctttggtta ggttgcgttt 3120
tggggtttgc ctttttttgt tgtcgcttaa atgcaatttg gttgtaaaga tttgattcct 3180
ttgtgttcat ctgttccgct tctcagcggt ccatctcagc gtctcccttc aggaaccgct 3240
gagtgtcctc tcttaacatc caagcctttt aatgaaatcg tactgaaatc tgtatcagct 3300
aagagtcctc caatcctggt cccattaact ccaagtgcct ttttgacagt gacaacagac 3360
agtccctcgc tttttgttgt tgttggtttt cttaacccct ttaatggaac tgcctggatt 3420
ttatacagtt attaa 3435

<210> 58
<211> 751
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504768CB1

<400> 58
agagctgtgt caccatgtgg gtcccggttg tcttcctcac cctgtccgtg acgtggattg 60
ggcagtctgc ggcggtgttc tggtgcaccc ccagtgggtc ctcacagctg cccactgcat 120
caggaacaaa agcgtgatct tgctgggtcg gcacagcctg tttcatcctg aagacacagg 180
ccaggtattt caggtcagcc acagcttccc acacccgctc tacgatatga gcctcctgaa 240
gaatcgattc ctcaggccag gtgatgactc cagccacgac ctcatgctgc tccgcctgtc 300
agagcctgcc gagctcacgg atgctgtgaa ggtcatggac ctgcccaccc aggagccagc 360
actggggacc acctgctacg cctcaggctg ggcagcatt gaaccagagg agttcttgac 420
cccaaagaaa cttcagtgtg tggacctcca tgttatttcc aatgacgtgt gtgcgcaagt 480
tcaccctcag aaggtgacca agttcatgct gtgtgctgga cgctggacag gggcaaaag 540
cacctgctcg ggtgattctg gggcccact tgtctgtaat ggtgtgcttc aaggtatcac 600
gtcatggggc agtgaaccat gtgccctgcc cgaaaggcct tccctgtaca ccaaggtggt 660
gcattaccgg aagtggatca aggacaccat cgtggccaac ccctgagcac ccctatacac 720
tccctattgt aaaaaacttg gcccttggaa a 751

<210> 59
<211> 1908
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500757CB1

<400> 59
ggccggctgg gcgtgcgctc gctccccgaa gccggggctg ggccggagcc gggcgagggc 60
tgggagctgg gccgggtccg gggacagcgg gcgaggggca gctgccggag ccgggcagcc 120
aggccgctca gggcagggga cagctggcgc cggttctgcg gtctccgggg cccagatgtg 180
aggcggcggc gccccccggcc cgagagcgca cgatgggggc cccgctcgcc gtagcgctgg 240

gcgccctcca ctacctggca cttttcctgc aactcggcgg cgccacgcgg cccgccggcc 300
acgcgccctg ggacaaccac gtctccggcc acgccctgtt cacagagaca ccccatgaca 360
tgacagcacg gacgggcgag gacgtggaga tggcctgctc cttccgcggc agcggctccc 420
cctcctactc gctggagatc cagtggtggt atgtacggag ccaccgggac tggaccgaca 480
agcaggcgtg ggcctcgaac caggtggtca aggtggtggg cagcaacatc tcccacaagc 540
tgcgcctgtc ccgggtgaag cccacggacg aaggcaccta cgagtgccgc gtcatcgact 600
tcagcgacgg caaggcccgg caccacaagg tcaaggccta cctgcgggtg cagccagggg 660
agaactccgt cctgcatctg cccgaagccc ctcccgccgc gcccgccccg ccgcccccca 720
agccaggcaa ggagctgagg aagcgctcgg tggaccagga ggcctgcagc ctctagactg 780
atgcccctgc ccccgcccat ccgcccccac gctgtacaga gtgcatgagg agccgccgga 840
ccaccgggga ccgactgcct gcgtccagcc gtgccccatc cccgaggccg cctgtggcca 900
ccatgtcggc cctctttcca ccaccccttg ctcagcatgt aagccccacc cacccctgcc 960
ctttcagacc cctgcggtga cctggctcgg agaaggtggc cctgggcacc aaggggccaa 1020
ccgccctgaa cactggggca gggaccatgc tggggcccgg ggccaccccc ttcctgtcac 1080
cagcttctgt ggagtccagt gttttgcttt gcttgcttgt cccccatcct gtcctgagcc 1140
ggggcccccc agcctcgcct ccctcctcct accatccctc acttggacct gggggtgtgg 1200
acagtgaccc ctccctgaat atggacttga atcttctgag cagaactagg gcctctcccc 1260
tggtgaagac ccagggaacc caggagggcc cttctggggc agtggctctg cagggtcact 1320
catggaggcc taggggaaca gcgagatgcc ccaccacctc ctggcgagtc cttcctgttc 1380
agctccctgt gcgaccctcc agggatgcag gggatccagg attctctgcc ctgtcacacg 1440
gcgagtcaga agggagggc ctttccctcg gacccatggc cccaggcaga gttttgcacc 1500
agcaggaccc ctttgagggc cttcaaggct ctcccaggag tccccctctg ccggcccccc 1560
aatgccccag ctccctcttg ggtcctgtgc caagtccgcc ccagggcctg gggctgttgg 1620
gagccaaggg ccccctggta ctcagttccc tcacgattcc cgatcacggg cacacctgcc 1680
ccctggttat ttgtaaatat ttctattgga cccaattctc ctcggaattg gctggcacct 1740
ctggctgcca cagctcagtg atgacgtggg ggaggtggga gaggccgagg gctttgccta 1800
ggggtgggtt gccctgtata catgatccag tctgtgacta ccagccaacc tgaataaagc 1860
ggttttaaaa aaaaaaaaaa aaaaaaaata tgcggcgcaa gcttattc 1908

<210> 60
<211> 1148
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1730616CB1

<400> 60
atccaatgac aagcggctcc ttgagaaagg gggggggggt actccttgaa agcattgtcc 60
aatgaaaatt accaacggca gacacggaga cagcgccggg gcggagggta cgatggagaa 120
tttcacggca ctgtttgggg ctcaggctga cccaccaccg cccccaaccg cactcggctt 180
cggaccagga aagcctccac ccccgccacc gcctcctgcg ggcgggggac ccggcacggc 240
cccgcctccc accgtggcca cggctcctcc tggcgcggac aagtcaggag ctggctgtgg 300
ccctttttac ctcatgaggg aactgccagg tagcacagag ctgacaggca gcacgaatct 360
gatcacacac tacaacttgg aacaagccta taataaattc tgtgggaaga aggtgaagga 420
gaagctaagt aacttcctgc ctgacctgcc agggatgatt gatctgcctg gttcccatga 480
taacagcagc ctccgctctc tcattgagaa gccccctatt ctcagtagct ctttcaatcc 540
tatcacaggg accatgctgg ccggcttccg cctccacact ggcccgttgc cggagcagtg 600
tcgtctgatg catattcagc ctcccaagaa gaagaataag cacaagcaca aacagagccg 660
tacccaggat cctgtccccc cagaaacacc atctgattca gatcacaaga agaagaaaaa 720
gaaaaaagaa gaggatcctg aacggaaaag gaagaagaaa gagaagaaga aaaagaagaa 780
tcgacatagt ccagaccacc caggtatggg cagctcccag gccagcagca gcagcagcct 840
acgctaatag gaccactgga ctctttgcca ggatggcttt tcctgctgta ctgaacctgc 900
tgataaaagc tgccttccag gctcttggac actgccttgg gagcatcctg cagctgggac 960
agaggccagc tcctgttggg ctcagttgag actaagtaaa ttaggagaga gagggacatg 1020

cttttgtagg ctcatcccag ttggttttct catggacatc tcttcctctc ccaggaagct 1080
tacaattttc ttctctctct tttgtgcaat ttgtctgatt taggacttgt tctgtgtttt 1140
ctttaaaa 1148

<210> 61
<211> 1776
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 190404CB1

<400> 61
agagagagag agagaactag tctcgagcta ttctcgagta tgtaatgatc aacagcatat 60
ttgaagcaat tttacagata ctttcccatc ccccaagtcg tagggagcat gggtatgatg 120
ctgtcgtcct cttggctttg ctggtgaact atagaaaata tgagtctgtg aatccttata 180
ttgtgaagct gtctatcgtg gatgatgagg ccacactcaa tggaatggga cttgtaattg 240
ctcaggcttt atctgagtac aacaggcagt ataaagacaa ggaagaagaa caccaaagtg 300
gtttttctc tgctttaaca aatatggtgg gcagcatgtt catagcagat gcccatgaga 360
aaatctcagt acaaaccaat gaggccattc ttctggcact ttatgaagct gttcatttaa 420
atcgcaactt catcacagta ttagctcaga gccatccaga aatgggcttg gtgacgaccc 480
ctgtcagtcc tgctcctaca accccagtca caccacttgg gaccacaccg ccttcctctg 540
atgttataag ttctgtggaa ctcccactgg atgcagatgt acagaccagt aatctactga 600
taaccttttt aaagtatagc tcaattgtca tgcaggacac caaagatgaa cacaggcttc 660
acagtggcaa actctgtctg attatcctta catgtattgc agaggatcaa tatgccaatg 720
catttttgca tgatgacaac atgaattttc gagtaaactt acatagaatg cctatgagac 780
acaggaagaa ggcagcagac aagaatcttc cctgccgtcc tttagtatgt gcagtactgg 840
acctgatggt agagtttatt gtaacacaca tgatgaagga gtttcctatg gatctctata 900
tacgctgcat ccaggtagta cacaaactgc tctgctacca gaagaagtgt cgggtacgcc 960
tgcattacac ctggcgggag ctctggtcag ccttgataaa tttgctgaag ttccttatgt 1020
caaatgagac tgtacttttg gccaaacaca acatttttac attagccctt atgattgtga 1080
acctatttaa tatgtttatc acatatggcg acacatttct gccaaccccc agcagctatg 1140
atgaacttta ctatgagatt atccgcatgc accagagctt tgacaacctc tactccatgg 1200
tcctgaggct ttctaccaat gcaggccagt ggaaggaagc agctagcaag gtgacccatg 1260
cattggttaa tatcagagcc atcatcaacc actttaaccc caaaattgag tcctacgctg 1320
ctgtgaatca catatcccaa ctgtcagagg agcaggtgct ggaggtggtg agagccaact 1380
atgacacgct cacgctgaag ctgcaggatg gcctggacca gtatgagcgc tactcagagc 1440
agcacaagga agctgccttc ttcaaagagc tggttcgatc cattagcacc aacgtccgga 1500
gaaacctggc cttccacaca ctcagccaag aagtcctgct caaggagttc tccactatct 1560
cctgaggcca cgcctacctg agcagcctct gactgccctt acccatgagg atcatgggct 1620
ggaggggggа gcgaggggag agggggctgc ccccgaggtt ggagagaaac acagatacca 1680
agttctcttg gtgaagaccg cacttcaatg gagcttgggc agcagggcag gaggtcaagg 1740
ggatccacta gttctaaacg ccccacccgc gtgcca 1776

<210> 62
<211> 1765
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500679CB1

<400> 62
ctctgagcag aactgacagc atgaaggtac ggggcccagg gtcgggggac tcatagcatg 60

ggggaactga gcccactcca gaggcccctg gccacagagg gcactatgaa ggcacaagga 120
gttctcttga aactcgcact cctggccctg ccgctgctgc tgcttctctc cacgcccccg 180
tgtgcccccc aggtctccgg gatccgagga gatggttttc cagaagagat tcaatggctc 240
agtaagtttc ttccgcggct ggaatgacta caagctgggc ttcggccgtg ctgatggaga 300
gtactggctg gggctgcaga acatgcacct cctgacactg aagcagaagt atgagctgcg 360
agtggacttg gaggactttg agaacaacac ggcctatgcc aagtacgctg acttctccat 420
ctccccgaac gcggtcagcg cagaggagga tggctacacc ctctttgtgg caggctttga 480
ggatggcggg gcaggtgact ccctgtccta ccacagtggc cagaagttct ctaccttcga 540
ccgggaccag gacctctttg tgcagaactg cgcagctctc tcctcaggag ccttctggtt 600
ccgcagctgc cactttgcca acctcaatgg cttctaccta ggtggctccc acctctctta 660
tgccaatggc atcaactggg cccagtggaa gggcttctac tactccctca aacgcactga 720
gatgaaaatc cgccggcct gaagggctgg ccccctcagg cacctttcct cccctggaca 780
cccatggtct ccatgagtgc tccctctgct gcccctgatg catgcttctg ctgattcccg 840
agcaccaact ccttacaagg gggccttgtg gctctcagcc atgccacatc cctgtcacac 900
acccagggca tccattccta agccagaccc ggctccccta cacctgaagt tacactgcca 960
gcagttcccc aggcctcttc cgagaggcac atggttctag cctggacctg gctgggctcc 1020
atgagaatga gttgcctcca ccctgtccca acagctgaca gccaggagcc actctcccag 1080
ctgcaggcct ttgtggtcca tcttgtcctg cttcctcact gtggacccct gtctgggcca 1140
ccctagtgtg ctaagctgag cagtgcagtg tgaacagggc ccatggtgta ttctaggcca 1200
cagcccagca ctcctctggg ctgctctcaa accatgtccc atcttcagca tccctcccac 1260
caacttactc ccctgtggtg agtaccgtgg aaccccagcc cacctcacta tcatactcag 1320
cttcccctga tggcccatcc cagcccctga agctctatgc caagaacaca gctaccgcac 1380
accaccctga aacagccaca gccaaggtag gcatgcatat gaggtcttcc ccataccctc 1440
tgggtgttga gaggtttagc cacatgaggg agcagaggac aatctctgca gggctgggag 1500
tgggtaggga ctgaaggtct caataaacct tcagaacctg aatgaactgg cttcatacac 1560
acaaacatat ttgtttatcc cccaaatgta ggcacctggc tcctccttgc tcccctgctg 1620
atggtgtcct accccgaact ccaaaaatta cacctggagt caggtgcaga agggaacctt 1680
gtatttcaca ggcctcattt tgatggcaaa aagacagtgt aataataaca taataataat 1740
aaaaatataa tactgaaaaa aaaaa 1765

<210> 63
<211> 1695
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500687CB1

<400> 63
gttgctccgg cggcgctcgg ggagggagcc agcagcctag ggcctaggcc cgggccacca 60
tggcgctgcc tccaggccca gccgccctcc ggcacacact gctgctcctg ccagcccttc 120
tgagctcagg cccctcccgg cacccatctc tgatatcaag tgactccaac aacctaaaac 180
tcaacaacgt gcgcctgcca cgggagaaca tgtccctccc gtccaacctt cagctcaatg 240
acctcactcc agattccaga gcagtgaaac cagcagaccg gcagatggct cagaacaaca 300
gccggccaga gcttctggac ccggagcccg gcggcctcct caccagccaa ggtttcatcc 360
gcctcccagt gctgggctat atctatcgag tgtccagcgt gagcagtgat gagatctggc 420
tctgagccga gggcgagaca ggagtattct cttggcctct ggacaccctc ccattcctcc 480
aaggcatcct ctacctagct aggtcaccaa cgtgaagaag ttatgccact gccacttttg 540
cttgccctcc tggctggggt gccctccatg tcatgcacgt gatgcatttc actgggctgt 600
aacccgcagg ggcacaggta tctttggcaa ggctaccagt tggacgtaag cccctcatgc 660
tgactcaggg tgggccctgc atgtgatgac tgggcccttc cagagggagc tctttggcca 720
ggggtgttca gatgtcatcc agcatccaag tgtggcatgg cctgctgtat accccacccc 780
agtactccac agcaccttgt acagtaggca tgggggcgtg cctgtgtggg ggacagggag 840
ggccctgcat ggatttcct ccttcctatg ctatgtagcc ttgttccctc aggtaaaatt 900
taggaccctg ctagctgtgc agaacccaat tgccctttgc acagaaacca acccctgacc 960

cagcggtacc ggccaagcac aaacgtcctt tttgctgcac acgtctctgc ccttcacttc 1020
ttctcttctg tccccacctc ctcttgggaa ttctaggtta cacgttggac cttctctact 1080
acttcactgg gcactagact tttctattgg cctgtgccat cgcccagtat tagcacaagt 1140
tagggaggaa gaggcaggcg atgagtctag tagcacccag gacggcttgt agctatgcat 1200
catttcccta cggcgttagc actttaagca catcccctag ggagggggt gagtgagggg 1260
cccagagccc tctttgtggc ttccccacgt ttggccttct gggattcact gtgagtgtcc 1320
tgagctctcg ggttgatgg tttttctctc agcatgtctc ctccaccacg ggaccccagc 1380
cctgaccaac ccatggttgc ctcatcagca ggaaggtgcc cttcctggag gatggtcgcc 1440
acaggcacat aattcaacag tgtggaagct ttaggggaac atggagaaag aaggagacca 1500
catacccсaa agtgacctaa gaacacttta aaaagcaaca tgtaaatgat tggaaattaa 1560
tatagtacag aatatatttt tcccttgttg agatcttctt ttgtaatgtt tttcatgtta 1620
ctgcctaggg cggtgctgag cacacagcaa gtttaataaa cttgactgaa ttcatttaca 1680
aaaaaaaaaa aaaaa 1695

<210> 64
<211> 2202
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500688CB1

<400> 64
gttgctccgg cggcgctcgg ggagggagcc agcagcctag ggcctaggcc cgggccacca 60
tggcgctgcc tccaggccca gccgccctcc ggcacacact gctgctcctg ccagcccttc 120
tgagctcagg ggggcctggc acccccagat tggcctggta tctggatgga cagctgcagg 180
aggccagcac ctcaagactg ctgagcgtgg gaggggaggc cttctctgga ggcaccagca 240
ccttcactgt cactgcccat cgggcccagc atgagctcaa ctgctctctg caggacccca 300
gaagtggccg atcagccaac gcctctgtca tccttaatgt gcaattcaag ccagagattg 360
cccaagtcgg cgccaagtac caggaagctc agggcccagg cctcctggtt gtcctgtttg 420
ccctggtgcg tgccaacccg ccggccaatg tcacctggat cgaccaggat gggccagtga 480
ctgtcaacac ctctgacttc ctggtgctgg atgcgcagaa ctacccctgg ctcaccaacc 540
acacggtgca gctgcagctc cgcagcctgg cacacaacct ctcggtggtg gccaccaatg 600
acgtgggtgt caccagtgcg tcgcttccag ccccaggccc ctcccggcac ccatctctga 660
tatcaagtga ctccaacaac ctaaaactca acaacgtgcg cctgccacgg gagaacatgt 720
ccctcccgtc caaccttcag ctcaatgacc tcactccaga ttccagagca gtgaaaccag 780
cagaccggca gatggctcag aacaacagcc ggccagagct tctggacccg gagcccggcg 840
gcctcctcac cagccaaggt ttcatccgcc tcccagtgct gggctatatc tatcgagtgt 900
ccagcgtgag cagtgatgag atctggctct gagccgaggg cgagacagga gtattctctt 960
ggcctctgga caccctccca ttcctccaag gcatcctcta cctagctagg tcaccaacgt 1020
gaagaagtta tgccactgcc acttttgctt gccctcctgg ctggggtgcc ctccatgtca 1080
tgcacgtgat gcatttcact gggctgtaac ccgcaggggc acaggtatct ttggcaaggc 1140
taccagttgg acgtaagccc ctcatgctga ctcagggtgg gccctgcatg tgatgactgg 1200
gcccttccag agggagctct ttggccaggg gtgttcagat gtcatccagc atccaagtgt 1260
ggcatggcct gctgtatacc ccaccccagt actccacagc accttgtaca gtaggcatgg 1320
gggcgtgcct gtgtggggga cagggagggc cctgcatgga ttttcctcct tcctatgcta 1380
tgtagccttg ttccctcagg taaaatttag gaccctgcta gctgtgcaga acccaattgc 1440
cctttgcaca gaaaccaacc cctgacccag cggtaccggc caagcacaaa cgtccttttt 1500
gctgcacacg tctctgccct tcacttcttc tcttctgtcc ccacctcctc ttgggaattc 1560
taggttacac gttggacctt ctctactact tcactgggca ctagactttt ctattggcct 1620
gtgccatcgc ccagtattag cacaagttag ggaggaagag gcaggcgatg agtctagtag 1680
cacccaggac ggcttgtagc tatgcatcat ttcctacgg cgttagcact ttaagcacat 1740
cccctagggg agggggtgag tgaggggccc agagccctct ttgtggcttc cccacgtttg 1800
gccttctggg attcactgtg agtgtcctga gctctcgggg ttgatggttt ttctctcagc 1860
atgtctcctc caccacggga ccccagccct gaccaaccca tggttgcctc atcagcagga 1920

aggtgccctt cctggaggat ggtcgccaca ggcacataat tcaacagtgt ggaagcttta 1980
ggggaacatg gagaaagaag gagaccacat accccaaagt gacctaagaa cactttaaaa 2040
agcaacatgt aaatgattgg aaattaatat agtacagaat atatttttcc cttgttgaga 2100
tcttcttttg taatgttttt catgttactg cctagggcgg tgctgagcac acagcaagtt 2160
taataaactt gactgaattc atttacaaaa aaaaaaaaaa aa 2202

<210> 65
<211> 779
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500697CB1

<400> 65
ctccctccct cgccttcctc cttcctagct cctctcctcc agggccagac tgagcccagg 60
ttgatttcag gcggacacca atagactcca cagcagctcc aggagcccag acaccggcgg 120
ccagaagcaa ggctaggagc tgctgcagcc atgtcggccc tcagcctcct cattctgggc 180
ctgctcacgg cagtgccacc tgccagctgt cagcaaggga gcctgcacac atgatcctga 240
ccgtcggaaa caaggcagat ggagtcctgg tgggaacaga tggaaggtac tcttcgatgg 300
cggccagttt caggtccagt gagcatgaga atgcctatga gaatgtgccc gaggaggaag 360
gcaaggtccg cagcaccccg atgtaacctt ctctgtggct ccaaccccaa gactcccagg 420
cacatgggat ggatgtccag tgctaccacc caagccccct ccttctttgt gtggaatctg 480
caatagtggg ctgactccct ccagccccat gccggcccta cccgcccttg aagtatagcc 540
agccaaggtt ggagctcaga ccgtgtctag gttggggctc ggctgtggcc ctggggtctc 600
ctgctcagct cagaagagcc ttctggagag gacagtcagc tgagcacctc ccatcctgct 660
cacacgtcct tccccataac tatggaaatg gccctaattt ctgtgaaata aagacttttt 720
gtatttctgg ggctgaggct cagcaacagc ccctcaggct tccagtgaaa aaaaaaaaa 779

<210> 66
<211> 3104
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500709CB1

<400> 66
aggacaaggt cgggagagga ggaggagggg agcgaggacg cagagaagga ggggaaagag 60
gacggagagg agcgaggacg gagggggggac gggaacgaag aggggcgagg acgttcagca 120
tccattgcga cagcgacaaa caggacaccg ccaggttccc ggcggaggca gcgcccaaga 180
tccccgttgc cgagcaaccg aagcatggcg gtggggcggg gtcggagcct gcgcatttcc 240
gcccccgccc gcccggctgg tcggaagtga cgccaggggg cggggccagc ggcgcggtcg 300
ggtgagaggc cgcggcggca ggtccacctg ggcttgcgaa ggcacagatt ccccgtccac 360
agctcacgac cagatgcacc agcaggagtc cacatcgagg acgtcctccg ggcactccca 420
cgaccagtga ccaggagtta aactttggga tgtgcccgtg atgttggacc acaaggactt 480
agaggccgaa atccacccct tgaaaaatga agaaagaaaa tcgcaggaaa atctgggaaa 540
tccatcaaaa aatgaggata acgtgaaaag cgcgcctcca cagtcccggc tctcccggtg 600
ccgagcggcg gcgtttttc tttcattgtt tctctgcctt tttgtggtgt tcgtcgtctc 660
attcgtcatc ccgtgtccag accggccggc gtcacagcga atgtggagga tagactacag 720
tgccgctgtt atctatgact ttctggctgt ggatgatata aacggggaca ggatccaaga 780
tgttcttttt ctttataaaa acaccaacag cagcaacaat ttcagccgat cctgtgtgga 840
cgaaggcttt tcctctccct gcacctttgc agctgctgtg tcgggggcca acggcagcac 900
gctctgggag agacctgtgg cccaagacgt ggccctcgtg gagtgtgctg tgccccagcc 960

aagaggcagt gaggcacctt ctgcctgcat cctggtgggc agacccagtt ctttcattgc 1020
agtcaacttg ttcacagggg aaaccctgtg gaaccacagc agcagcttca gcgggaatgc 1080
gtccatcctg agccctctgc tgcaggtgcc tgatgtggac ggcgatgggg ccccagacct 1140
gctggttctc acccaggagc gggaggaggt tagtggccac ctctactccg gcagcaccgg 1200
gcaccagatt ggcctcagag gcagccttgg tgtggacggg gaaagtggct tcctccttca 1260
cgtcaccagg acaggtgccc actacatcct ctttccctgc gcaagctccc tctgcggctg 1320
ctctgtgaag ggtctctacg agaaggtgac cgagagcggc ggcccgttca agagtgaccc 1380
gcactgggag agcatgctca atgccaccac ccgcaggatg ctttcccaca gctctggagc 1440
agtgcgctac ctgatgcatg tcccagggaa cgccggtgca gatgtgcttc ttgtgggctc 1500
agaggccttc gtgctgctgg acgggcagga gctgacgcct cgctggacac ccaaggcagc 1560
ccatgtcctg agaaaaccca tcttcggccg ctacaaacca gacaccttgg ctgtagccgt 1620
tgaaaacgga actggcaccg acagacagga gaccggggag gcccggcaca gcctgtacat 1680
gttccacccc accctgccgc gcgtgctgct ggagctggcc aatgtctcta cccacattgt 1740
cgcctttgac gccgtcctgt ttgagccaag ccgccacgcc gcctacatcc ttctgacagg 1800
cccggcagac tcagaggcac ccggcctggt ctctgtgatc aagcacaagg tgcgggacct 1860
tgtcccaagc agcagggtgg tccgcctggg tgagggtggg ccagacagtg accaagccat 1920
cagggaccgg ttctcccggc tgcggtacca gagtgaggcg tagaggcacg ccagccagag 1980
cctgtggaga gactccgcct gctgacacta aacgtcctgg gaagtgggcc cttccctggg 2040
tctctgcact gactccccca ctcctgaccc tggtgatggt cgccactggg cagcagcagc 2100
cttaccagtc ctccatgatc acacccaggg acctgcatgg gtgaggggac accctgggcc 2160
tctctcccgc ccagcatcct ccctgagtcc ccacacaggg cctcactctg caccccacca 2220
gggtcccgct cacaccaggc agccttcata gtggtctccc tggccacctt gggcagagct 2280
gggtcatgca gcaccccatc cttacccggt gccctctcct tgccagcttc tccccaggcc 2340
agagcggcca tcgcgtagaa agaaccaggg tgtccccggg acaggccgtc ccccaccccca 2400
tcctgtagaa gtccattccc cttttccctc ctgtgctctg tcccccaagg agtcatggaa 2460
ctcagggtac tgggcctcaa cgggaacctg agacagctcc agcttcgcag cccttcccgg 2520
agctacaggg ggatcctcta gcatgggggg tgtgacttgg ttcctttgac caggtcctgt 2580
gaggaagcct ggagcaaggg tctcccccag caggatgggt ggggcctgct ctggagctga 2640
gcccgtggcc gctcacaggt gtccttagtg gtgttgcagc tgtctactgg ctgcatgtgc 2700
tgtgaatatc ccaaggaact ggctgtggaa tgcgtgtttg ggtcagtctg tgccctctca 2760
gtagacactg gagctgctct gtccctgaag aggccccgtg ccccaggcat ggcaagcgcc 2820
tgcctctccc cttccggtgc tcacacgccc acgccgtgcc acccgatgca ggactcacct 2880
ctgtgccttg ctgctcctga ggcccaaggg cagccatggt gctctgtact gctcgggccg 2940
cccaggtcac agagcctgag cttcgtagcc aaagcagcct gatgacccac ccaccaagga 3000
agaaagcaga ataaacattt ttgcactgcc tgaaaaaccc cggtggtcag gcgtgagcct 3060
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaatggtgc ggtc 3104

<210> 67
<211> 3233
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500711CB1

<400> 67
aggacaaggt cgggagagga ggaggagggg agcgaggacg cagagaagga ggggaaagag 60
gacggagagg agcgaggacg gaggggggac gggaacgaag aggggcgagg acgttcagca 120
tccattgcga cagcgacaaa caggacaccg ccaggttccc ggcggaggca gcgcccaaga 180
tccccgttgc cgagcaaccg aagcatggcg gtggggcggg gtcggagcct gcgcatttcc 240
gcccccgccc gcccggctgg tcggaagtga cgccaggggg cggggccagc ggcgcggtcg 300
ggtgagaggc cgcggcggca ggtccacctg ggcttgcgaa ggcacagatt ccccgtccac 360
agctcacgac cagatgcacc agcaggagtc cacatcgagg acgtcctccg ggcactccca 420
cgaccagtga ccaggagtta aactttggga tgtgcccgtg atgttggacc acaaggactt 480
agaggccgaa atccacccct tgaaaaatga agaaagaaaa tcgcaggaaa atctgggaaa 540

tccatcaaaa aatgaggata acgtgaaaag cgcgcctcca cagtcccggc tctcccggtg 600
ccgagcggcg gcgtttttc tttcattgtt tctctgcctt tttgtggtgt tcgtcgtctc 660
attcgtcatc ccgtgtccag accggccggc gtcacagcga atgtggagga tagactacag 720
tgccgctgtt atctatgact ttctggctgt ggatgatata aacggggaca ggatccaaga 780
tgttcttttt ctttataaaa acaccaacag cagcaacaat ttcagccgat cctgtgtgga 840
cgaagctgct gtgtcggggg ccaacggcag cacgctctgg gagagacctg tggcccaaga 900
cgtggccctc gtggagtgtg ctgtgcccca gccaagaggc agtgaggcac cttctgcctg 960
catcctggtg ggcagaccca gttctttcat tgcagtcaac ttgttcacag ggaaaccct 1020
gtggaaccac agcagcagct tcagcgggaa tgcgtccatc ctgagccctc tgctgcaggt 1080
gcctgatgtg gacggcgatg ggccccaga cctgctggtt ctcacccagg agcgggagga 1140
ggttagtggc cacctctact ccggcagcac cgggcaccag attggcctca gaggcagcct 1200
tggtgtggac ggggaaagtg gcttcctcct tcacgtcacc aggacaggtg cccactacat 1260
cctctttccc tgcgcaagct ccctctgcgg ctgctctgtg aagggtctct acgagaaggt 1320
gaccgggagc ggcggcccgt tcaagagtga cccgcactgg gagagcatgc tcaatgccac 1380
cacccgcagg atgctttccc acagctctgg agcagtgcgc tacctgatgc atgtcccagg 1440
gaacgccggt gcagatgtgc ttcttgtggg ctcagaggcc ttcgtgctgc tggacgggca 1500
ggagctgacg cctcgctgga cacccaaggc agcccatgtc ctgagaaaac ccatcttcgg 1560
ccgctacaaa ccagacacct tggctgtagc cgttgaaaac ggaactggca ccgacagaca 1620
gatcctgttt ctggaccttg gcactggagc cgtcctgtgt agcctagccc tcccgagcct 1680
ccctgggggt ccactgtccg ccagcctgcc gaccgcagac caccgctcag ccttcttctt 1740
ctggggcctc cacgagctgg ggagcaccag cgagacggag accggggagg cccggcacag 1800
cctgtacatg ttccacccca ccctgccgcg cgtgctgctg gagctggcca atgtctctac 1860
ccacattgtc gcctttgacg ccgtcctgtt tgagccaagc cgccacgccg cctacatcct 1920
tctgacaggc ccggcagact cagaggcacc cggcctggtc tctgtgatca agcacaaggt 1980
gcgggacctt gtcccaagca gcagggtggt ccgcctgggt gagggtgggc cagacagtga 2040
ccaagccatc agggaccggt tctcccggct gcggtaccag agtgaggcgt agaggcacgc 2100
cagccagagc ctgtggagag actccgcctg ctgacactaa acgtcctggg aagtgggccc 2160
ttccctgggt ctctgcactg actcccccac tcctgaccct ggtgatggtc gccactgggc 2220
agcagcagcc ttaccagtcc tccatgatca cacccaggga cctgcatggg tgaggggaca 2280
ccctgggcct ctctcccgcc cagcatcctc cctgagtccc cacacagggc ctcactctgc 2340
accccaccag ggtcccgctc acaccaggca gccttcatag tggtctccct ggccaccttg 2400
ggcagagctg ggtcatgcag caccccatcc ttacccggtg ccctctcctt gccagcttct 2460
ccccaggcca gagcggccat cgcgtagaaa gaaccagggt gtccccggga caggccgtcc 2520
cccaccccat cctgtagaag tccattcccc ttttccctcc tgtgctctgt cccccaagga 2580
gtcatggaac tcagggtact gggcctcaac gggaacctga gacagctcca gcttcgcagc 2640
ccttcccgga gctacagggg gatcctctag catggggggt gtgacttggt tcctttgacc 2700
aggtcctgtg aggaagcctg gagcaagggt ctcccccagc aggatgggtg gggcctgctc 2760
tggagctgag cccgtggccg ctcacaggtg tccttagtgg tgttgcagct gtctactggc 2820
tgcatgtgct gtgaatatcc caaggaactg gctgtggaat gcgtgtttgg gtcagtctgt 2880
gccctctcag tagacactgg agctgctctg tccctgaaga ggccccgtgc cccaggcatg 2940
gcaagcgcct gcctctcccc ttccggtgct cacacgccca cgccgtgcca cccgatgcag 3000
gactcacctc tgtgccttgc tgctcctgag gcccaagggc agccatggtg ctctgtactg 3060
ctcgggccgc ccaggtcaca gagcctgagc ttcgtagcca aagcagcctg atgacccacc 3120
caccaaggaa gaaagcagaa taaacatttt tgcactgcct gaaaaacccc ggtggtcagg 3180
cgtgagccta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaatggtgcg gtc 3233

<210> 68
<211> 1808
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500723CB1

<400> 68

ggagcacggc gctggggccg cccgcagcgc tcactcgctc gcactcagtc gcgggaggct 60
tccccgcgcc ggccgcgtcc cgcccgctcc ccggcaccag aagttcctct gcgcgtccga 120
cggcgacatg ggcgtcccca cggccccgga ggccggcagc tggcgctggg gatccctgct 180
cttcgctctc ttcctggctg cgtccctagg accttcacct gcaccatgga ggccaccagg 240
ctgccaacac cagccacgac ctggctcagc gccacgggct ggagtcggcc tccgaccacc 300
atggcaactt ctccatcacc atgcgcaacc tgaccctgct ggatagcggc ctctactgct 360
gcctggtggt ggagatcagg caccaccact cggagcacag ggtccatggt gccatggagc 420
tgcaggtgca gacaggcaaa gatgcaccat ccaactgtgt ggtgtaccca tcctcctccc 480
aggagagtga aaacatcacg gctgcagccc tggctacggg tgcctgcatc gtaggaatcc 540
tctgcctccc cctcatcctg ctcctggtct acaagcaaag gcaggcagcc tccaaccgcc 600
gtgcccagga gctggtgcgg atggacagca acattcaagg gattgaaaac cccggctttg 660
aagcctcacc acctgcccag ggatacccg aggccaaagt caggcacccc ctgtcctatg 720
tggcccagcg gcagccttct gagtctgggc ggcatctgct ttcggagccc agcacccccc 780
tgtctcctcc aggccccgga gacgtcttct tcccatccct ggaccctgtc cctgactctc 840
caaactttga ggtcatctag cccagctggg ggacagtggg ctgttgtggc tgggtctggg 900
gcaggtgcat ttgagccagg gctggctctg tgagtggcct ccttggcctc ggccctggtt 960
ccctccctcc tgctctgggc tcagatactg tgacatccca gaagcccagc ccctcaaccc 1020
ctctggatgc tacatgggga tgctggacgg ctcagcccct gttccaagga ttttggggtg 1080
ctgagattct cccctagaga cctgaaattc accagctaca gatgccaaat gacttacatc 1140
ttaagaagtc tcagaacgtc cagcccttca gcagctctcg ttctgagaca tgagccttgg 1200
gatgtggcag catcagtggg acaagatgga cactgggcca ccctcccagg caccagacac 1260
agggcacggt ggagagactt ctcccccgtg gccgccttgg ctcccccgtt ttgcccgagg 1320
ctgctcttct gtcagacttc ctctttgtac cacagtggct ctggggccag gcctgcctgc 1380
ccactggcca tcgccacctt ccccagctgc ctcctaccag cagtttctct gaagatctgt 1440
caacaggtta agtcaatctg ggcttccac tgcctgcatt ccagtcccca gagcttggtg 1500
gtcccgaaac gggaagtaca tattggggca tggtggcctc cgtgagcaaa tggtgtcttg 1560
ggcaatctga ggccaggaca gatgttgccc cacccactgg agatggtgct gagggaggtg 1620
ggtggggcct tctgggaagg tgagtggaga ggggcacctg ccccccgccc tccccatccc 1680
ctactcccac tgctcagcgc gggccattgc aagggtgcca cacaatgtct tgtccaccct 1740
gggacacttc tgagtatgaa gcgggatgct attaaaaact acatggggaa acaggtgcaa 1800
aaaaaaa 1808

<210> 69
<211> 1672
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500764CB1

<400> 69
cggccgcacc cggccggagc ggagggcaga gcgcgcgccc agttgcccgg gcaccaaatc 60
ggagcgcggc gtgcgggagg gcccagagca ggactggaaa tgtcctggcc gcgccgcctc 120
ctgctcagat acctgttccc ggccctcctg cttcacgctg tgaagtaagg tcaggcccag 180
agttcatcac aaggtcctac agattctacc acaataacac cttcaaggcc taccaatttt 240
attatggcag caaccggtgc acaaatccca cttatactct catcatccgg ggcaagatcc 300
gcctccgcca ggcctcctgg atcatccgag gggcacgga agccgactac cagctgcaca 360
acgtccaggt gatctgccac acagaggcgg tggccgagaa gctgggccag caggtgaacc 420
gcacatgccc gggcttcctc gcagacgggg gtccctgggt gcaggacgtg gcctatgacc 480
tctggcgaga ggagaacggc tgtgagtgca ccaaggccgt gaactttgcc atgcatgaac 540
ttcagctcat ccgggtggag aagcagtacc ttcaccacaa cctcgaccac ctggtcgagg 600
agctcttcct tggtgacatt cacactgatg ccacccagag gatgttctac cggccctcca 660
gttaccagcc ccctctgcag aatgccaaga accacgacca tgcctgcatc gcctgtcgga 720
tcatctatcg gtcagacgag caccaccctc ccatcctgcc cccaaaggca gacctgacca 780
tcggcctgca cggggagtgg gtgagccagc gctgtgaggt gcgccccgaa gtcctcttcc 840

tcacccgcca cttcatcttc catgacaaca acaacacctg ggagggccac tactaccact 900
actcagaccc ggtgtgcaag caccccacct tctccatcta cgcccggggc cgctacagcc 960
gcggcgtcct ctcgtccagg gtcatgggag gcaccgagtt cgtgttcaaa gtgaatcaca 1020
tgaaggtcac ccccatggat gcggccacag cctcactgct caacgtcttc aacgggaatg 1080
agtgcggggc cgagggctcc tggcaggtgg gcatccagca ggatgtgacc cacaccaatg 1140
gctgcgtggc cctgggcatc aaactacctc acacggagta cgagatcttc aaaatggaac 1200
aggatgcccg gggcgctat ctgctgttca acggtcagag gcccagcgac gggtccagcc 1260
cagacaggcc agagaagaga gccacgtcct accagatgcc cttggtccag tgtgcctcct 1320
cttcgccgag ggcagaggac ctcgcagaag acagtggaag cagcctgtat ggccgggccc 1380
ctgggaggca cacctggtcc ctgctgctgg ctgcacttgc ctgccttgtc cctctgctgc 1440
attggaacat ccgcagatag aagttttaga aagttctatt tttccaaacc aggattcctt 1500
actattgaca gatttctttt accaaaagaa aagacattta ttcttttgat gcacttgaat 1560
gccagagaac tgtccttctt tttctcctct ccctccctcc cagcccctga gtcatgaaca 1620
gcaaggagtg tttgaagttt ctgctttgaa ctccgtccag cctgatccct gg 1672

<210> 70
<211> 1347
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500772CB1

<400> 70
cactgcagcg caggtgagct ctcctgagga cctctctgtc agctcccctg attgtaggga 60
ggatccagtg tggcaagaaa ctcctccagc ccagcaagca gctcaggatg ttcctgaagg 120
ccgtggtcct gaccctggcc ctggtggctg tcgccgtgcc ctcttccagg acaaacttgg 180
agaagtgaac acttacgcag gtgacctgca gaagaagctg gtgccctttg ccaccgagct 240
gcatgaacgc ctggccaagg actcggagaa actgaaggag gagattggga aggagctgga 300
ggagctgagg gcccggctgc tgccccatgc caatgaggtg agccagaaga tcggggacaa 360
cctgcgagag cttcagcagc gcctggagcc ctacgcggac cagctgcgca cccaggtcaa 420
cacgcaggcc gagcagctgc ggcgccagct gacccccctac gcacagcgca tggagagagt 480
gctgcgggag aacgccgaca gcctgcaggc ctcgctgagg ccccacgccg acgagctcaa 540
ggccaagatc gaccagaacg tggaggagct caagggacgc cttacgccct acgctgacga 600
attcaaagtc aagattgacc agaccgtgga ggagctgcgc cgcagcctgg ctccctatgc 660
tcaggacacg caggagaagc tcaaccacca gcttgagggc ctgaccttcc agatgaagaa 720
gaacgccgag gagctcaagg ccaggatctc ggccagtgcc gaggagctgc ggcagaggct 780
ggcgcccttg gccgaggacg tgcgtggcaa cctgaggggc aacaccgagg ggctgcagaa 840
gtcactggca gagctgggtg ggcacctgga ccagcaggtg gaggagttcc gacgccggt 900
ggagccctac ggggaaaact tcaacaaagc cctggtgcag cagatggaac agctcaggca 960
gaaactgggc ccccatgcgg gggacgtgga aggccacttg agcttcctgg agaaggacct 1020
gagggacaag gtcaactcct tcttcagcac cttcaaggag aaagagagcc aggacaagtc 1080
tctctccctc cctgagctgg agcaacagca ggaacagcag caggagcagc agcaggagca 1140
ggtgcagatg ctggcccctt tggagagctg agctgcccct ggtgcactgg ccccaccctc 1200
gtggacacct gccctgccct gccacctgtc tgtctgtctg tcccaaagaa gttctggtat 1260
gaacttgagg acacatgtcc agtgggaggt gagaccacct ctcaatattc aataaagctg 1320
ctgagaatct agcctcaaaa aaaaaaa 1347

<210> 71
<211> 729
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7501350CB1

<400> 71
gctggtgctg gtacaggttg gggtcgctga tgaacgcgaa gatgatgcag gcaacgaagg 60
tctccagcac cttcagcagc ccgggtacgg tggccatata gccagtgatc tcgccgggcg 120
gcgtggcagg aagaggggga ctctgtggtc agggaacttc tcgctgagca cagctgcaca 180
gtgctgtcag aacggccgat ctccagccca agatgattcc agcagtggtc ttgctcttac 240
tcctttggt tgaacaagca gatccaagtg cgaaaggcag ctataaccag ctatgagaaa 300
tcagatggtg tttacacggg cctgagcacc aggaaccagg agacttacga gactctgaag 360
catgagaaac caccacagta gctttagaat agatgcggtc atattcttct ttggcttctg 420
gttcttccag ccctcatggt tggcatcaca tatgcctgca tgccattaac accagctggc 480
cctaccccta taatgatcct gtgtcctaaa ttaatataca ccagtggttc ctcctccctg 540
ttaaagacta atgctcagat gctgtttacg gatatttata ttctagtctc actctcttgt 600
cccacccttc ttctcttccc cattcccaac tccagctaaa atatgggaag ggagaacccc 660
caataaaact gccatggact ggaaaaaaaa aaaaaaaaaa ttcctctggt cgaacttttc 720
ccttacctt 729

<210> 72
<211> 1188
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506396CB1

<220>
<221> unsure
<222> (1) ... (1188)
<223> a, t, c, g, or other

<400> 72
aggactgctt tatggacttg catcaggagt agcaacagga ctctggactg tcttcacact 60
aggtggtgtg ggcagcaaag cggctgnctc tccagagctt tcttcccctc tggccaacca 120
gagnctcctg cttctgctgg tgttggccaa tctgacagat gcctcaaatg cgccaaaccc 180
ctacagacaa gccattatgt ccttcaagaa cacacaaggt tttaccaatt cttgagattc 240
tgtatcatgt tgaagaaagg aattcacacc atgtgtatat ggcccttata atattgttga 300
tccttacgga agatgatggc ttcaacagat ccattcatga agtgatacta aaaaatatta 360
cttggtattc agaacgagtt ttaactgaaa tctccttggg gagtctcctg atcctggtgg 420
taataagaac cattcaatac aacatgacta ggacacgaga caagtacctt cacacaaatt 480
gtttggcagc tttagcaaat atgtcggcac agtttcgttc tctccatcag tatgctgccc 540
agaggatcat cagtttattt tctttgctgt ctaaaaaaca caacaaagtt ctggaacaag 600
ccacacagtc cttgagaggt tcgctgagtt ctaatgatgt tcctctacca gattatgtga 660
tctccttctt tagctcaagg ttgctgcaag ctggagctga gctgtcagtg gaacgggtcc 720
tggaaatcat taagcaaggc gtcgttgcgc tgcccaaaga cagactgaag aaatttccag 780
aattgaaatt caaatatgtg gaagaggagc agcccgagga gtttttttatc ccctatgtct 840
ggtctcttgt ctacaactca gcagtcggcc tgtactggaa tccacaggac atccagctgt 900
tcaccatgga ttccgactga gggcaggatg ctctcccacc cggacccctc cagccaagca 960
gcccttcaag ttctttttatt tctgggtaac agaagtagac agacaggtta cttggtgtat 1020
cttctgttaa agaggattgc acgagtgtgt tttcctcaca cactttgatt tggagaattg 1080
gtgctagttg gcaatagata actcagcgta gatagtattg caaaaagggg aggaaataca 1140
caacaataat aaatgtaaaa acctgctatt caaaaaaaaa aaaaaaaa 1188

<210> 73
<211> 1886
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505917CB1

<400> 73
gggggcacct ctggtgacca agaccgggct gcgctccaaa gaggccgttg ggcctggagt 60
ggggttgggg gggtccgaga ggagttgggt gacatccccc accccatccc gggtccagct 120
gtttcagccc ctctcggcgc gccgatacta ttagccccac ccgtcctcca tcgagtcccg 180
tgccgctccc aaaccgcacg ataagcccca cagggagtgc gccataggcc ggggcgcgtc 240
acggggccgg ggcggggcgg agtccggacg tcgggagcag gatggcggcg gagcaggacc 300
ccgaggcgcg cgcggcggcg cggccgctgc tcactgacct ctaccaggcc accatggcgt 360
tgggctattg gcgcgcgggc cgggcgcggg acgccgccga gttcgagctc ttcttccgcc 420
gctgcccgtt cggcggcgcc ttcgccttgg ccgccggctt gcgcgactgt gtgcgcttcc 480
tgcgcgcctt ccgcctgcgg gacgccgacg tgcagttcct ggcctcggtg ctgcccccag 540
acacggatcc tgcgttcttc gagcaccttc gggccctcga ctgctccgag gtgacggtgc 600
gagccctgcc cgagggctcc ctcgccttcc ccggagtgcc gctcctgcag gtgtccgggc 660
cgctcctggt ggtgcagctg ctggagacac cgctgctctg cctggtcagc tacgccagcc 720
tggtggccac caacgcagcg cggcttcgct tgatcgcagg gccagagaag cggctgctag 780
agatgggcct gaggcgggct cagggccccg atgggggcct gacagcctcc acctacagct 840
acctgggcgg cttcgacagc agcagcaacg tgctagcggg ccagctgcga ggtgtgccgg 900
tggccgggac cctggcccac tccttcgtca cttccttttc aggcagcgag gtgccccctg 960
acccgatgtt ggcgccagca gctggtgagg gccctggggt ggacctggcg gccaaagccc 1020
aggtgtggct ggagcaggtg tgtgcccacc tggggctggg ggtgcaggag ccgcatccag 1080
gcgagcgggc agcctttgtg gcctatgcct tggcttttcc ccgggccttc cagggcctcc 1140
tggacaccta cagcgtgtgg aggagtggtc tccccaactt cctagcagtc gccttggccc 1200
tgggagagct gggctaccgg gcagtgggcg tgaggctgga cagtggtgac ctgctacagc 1260
aggctcagga gatccgcaag gtcttccgag ctgctgcagc ccagttccag gtgccctggc 1320
tggagtcagt cctcatcgta gtcagcaaca acattgacga ggaggcgctg gcccgactgg 1380
cccaggagct ggtggccgtg gggggccagc cacgaatgaa gctgaccgag gaccccgaga 1440
agcagacgtt gcctgggagc aaggctgctt tccggctcct gggctctgac gggtctccac 1500
tcatggacat gctgcagtta gcagaagagc cagtgccaca ggctgggcag gagctgaggg 1560
tgtggcctcc aggggcccag gagccctgca ccgtgaggcc agcccaggtg gagccactac 1620
tgcggctctg cctccagcag ggacagctgt gtgagccgct cccatccctg gcagagtcta 1680
gagccttggc ccagctgtcc ctgagccgac tcagccctga gcacaggcgg ctgcggagcc 1740
ctgcacagta ccaggtggtg ctgtccgaga ggctgcaggc cctggtgaac agtctgtgtg 1800
cggggcagtc cccctgagac tcggagcggg gctgactgga aacaacacga atcactcact 1860
tttccccaca aaaaaaaaaa aaaaaa 1886

<210> 74
<211> 1067
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500701CB1

<400> 74
ctcgagccga attcggctcg agcggcgctc ctgcctccct gcagggagct gcttatggga 60
caccgcttcc tgcgcggcct cttaacgctg ctgctgccgc cgccacccct gtatacccgg 120
caccgcatgc tcggtccaga gtccgtcccg cccccaaaac gatcccgcag caaactcatg 180
gcaccgcccc gaatcgggac gcacaatggc accttccact gcgacgaggc actggcatgc 240
gcactgcttc gcctcctgcc ggagtaccgg gatgcagaga ttgtgcggac ccgggatccc 300
gaaaaactcg cttcctgtga catcgtggtg gacgtggggg gcgagtacga ccctcggaga 360

caccgatatg accatcacca gaggcagggt tcaagcgtgc aatggatctg gttcaagagg 420
agtttctgca gagattagat ttctaccaac acagctggct gccagcccgg gccttggtgg 480
aagaggccct tgcccagcga ttccaggtgg acccaagtgg agagattgtg gaactggcga 540
aaggtgcatg tccctggaag gagcatctct accacctgga atctgggctg tcccctccag 600
tggccatctt ctttgttatc tacactgacc aggctggaca gtggcgaata cagtgtgtgc 660
ccaaggagcc ccactcattc caaagccggc tgcccctgcc agagccatgg cggggtcttc 720
gggacgaggc cctggaccag gtcagtggga tccctggctg catcttcgtc catgcaagcg 780
gcttcattgg cggtcaccgc acccgagagg gtgccttgag catggcccgt gccaccttgg 840
cccagcgctc atacctccca caaatctcct agtctaataa aaccttccat ctcatactga 900
caaaaaaaaa atggtagctc ggggggcgtt aggatagatc tttagacccg gtggggtttc 960
taagctggag aagtgttaga aaaaaagggc ggccgcccat caaattagtt ctctgccccg 1020
ggattttttt tcgggccggt acccttgggg ggaccaaatt tcccctt 1067

<210> 75
<211> 1220
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500702CB1

<400> 75
ctcgagccga attcggctcg agcggcgctc ctgcctccct gcagggagct gcttatggga 60
caccgcttcc tgcgcggcct cttaacgctg ctgctgccgc cgccacccct gtatacccgg 120
caccgcatgc tcggtccaga gtccgtcccg cccccaaaac gatcccgcag caaactcatg 180
gcaccgcccc gaatcgggac gcacaatggc accttccact gcgacgaggc actggcatgc 240
gcactgcttc gcctcctgcc ggagtaccgg gatgcagaga ttgtgcggac ccgggatccc 300
gaaaaactcg cttcctgtga catcgtggtg gacgtggggg gcgagtacga ccctcggaga 360
caccgatatg accatcacca gagatgtatg agaactttgt ggaggaggtg gatgctgtgg 420
acaatgggat ctcccagtgg gcagaggggg agcctcgata tgcactgacc actaccctga 480
gtgcacgagt tgctcgactt aatcctacct ggaaccaccc cgaccaagac actgaggcag 540
ggttcaagcg tgcaatggat ctggttcaag aggagtttct gcagagatta gatttctacc 600
aacacagctg gctgccagcc cgggccttgg tggaagaggc ccttgcccag cgattccagg 660
tggacccaag tggagagatt gtggaactgg cgaaaggtgc atgtccctgg aaggagcatc 720
tctaccacct ggaatctggg ctgtcccctc cagtggccat cttctttgtt atctacactg 780
accaggctgg acagtggcga atacagtgtg tgcccaagga gccccactca ttccaaagcc 840
ggctgcccct gccagagcca tggcggggtc ttcgggacga ggccctggac caggtcagtg 900
ggatccctgg ctgcatcttc gtccatgcaa gcggcttcat tggcggtcac cgcacccgag 960
agggtgcctt gagcatggcc cgtgccacct tggcccagcg ctcatacctc ccacaaatct 1020
cctagtctaa taaaaccttc catctcatac tgacaaaaaa aaaatggtag ctcggggggc 1080
gttaggatag atctttagac ccggtggggt ttctaagctg gagaagtgtt agaaaaaaag 1140
ggcggccgcc catcaaatta gttctctgcc ccgggatttt ttttcgggcc ggtacccttg 1200
gggggaccaa atttcccctt 1220

<210> 76
<211> 3280
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6044343CB1

<400> 76
gcggcggcgc ctggatgccg gcccggggct gggccggaga ctggcgcccg ccccgaagcg 60

actttcggtc cccgacccgc tccgcctgac ccctatgatg aagagggcgt ccactggagt 120
tcttgctgtc gacatctttg acaagttcaa cttcacggga gccagagtcc catggttcga 180
cgccattcat gaagaattcc ccagggagct ggagtcctcc atctccttcc cagccaactt 240
cttcaaacca cctgaagaaa aagaaggccc cctggtaagg ccagccagcc ggaggaccac 300
cccgcagacc acacgcccag gcctggcacc gagggggagg ccccgatcag caggtgaaag 360
cgacaccctg atgacactgg ctagttcacc ttcactctgg tcatcgttta ctgtaccccg 420
gggcagctcc cgcctgagcc ctatgacccc aaccgccgca gcctctggct gcctcactgg 480
cccatcatta atacgtcgat ggtgagcgca ctggtgtaca gcgagggggc tccacttccg 540
agacccctgg agaggcccgt cctggtggag ttcaccctgc tggagatgga gaagcgaacc 600
aagcctgtct gtgtgttctg gaaccattcc ctggcgtaaa tgtgatgccg ggagttctgc 660
atcttcatgt gagggctttg ccaagacaca gagaacccgc agggttgttg gctcagcggc 720
agatcgtccc acggggacag tctcttgtgg tctgtgtctc cccctcagc cccctgtggg 780
gcagtgagcc tcctgctctg cttctgtggg tgcatcagag ccccccaga tgcccacacg 840
cagacgcgca caggcctgca cgtggaaaca cagacatgcg cacacgcagg cccacacaca 900
gacgcacaca gagctgcacg tggagacacg tgcgcacacg caggcccaca cgcaggccca 960
cacacaggtg cacgcggacc tgcacgtgga gacacagatg tgtgcacatg caggcccaca 1020
cacaggtgag cgtggacctg tgtgaacaga tgcgcgcaga cctgcacgtg gagacacaga 1080
catgcaggcg cacaaagaga tccctcaagg ctgacaccca ggaaggggcc tctcacttgt 1140
gggactcccg agatgcagtg gccaacacta gtgtagtgcc cagttcactg ctggataaga 1200
cagaggcctg ttctggtcac acatgggagg cagaacccag agactgggcc caggagccct 1260
tctgctgaca ctgggaactc ccagctacgt gtgggggtcc ccataccaga caaaggtccc 1320
tgaccttagt cttgcccgag aggccgacac agcccagctt tggggtctgg ctttacccac 1380
aagaggccac accttgccac agcactgttt atctggcctg tttcagaagc accgtcagag 1440
tggcgaaggc aggaggtggt gcacgagagt ctacgttcta gcatccattc aagtgaggga 1500
aaggcgtttc gtacttagaa aaatgcaaaa ttaatgattc tcacccacac atgagactgt 1560
ggattcaccc atgtatgaga cagtggattc agagccatgc tgtggctctc atctgctggg 1620
agccagtagt gtttgtgctt tggttttgct tttttagcta catggtttat gtgaggtgcg 1680
gacttacttg agttgctctg tctttacagg gagcttgctt tttctctgct ggtttgtttt 1740
cacttttgct gtctctgttt agaagaaaac attagaactg ggtgtggtg gaaggatggg 1800
agcacaatct gggaggggaa gacctgctct ctctccacat accaggaggg agtggcagga 1860
aggggagaca ctgtgccgga ctcagggagt caagagcgtg agcaccccca ggcttcaatc 1920
tgctgaacgt tgtgcatctg gggcgcccg tgcagcagga acactctccc catcacaatt 1980
cctctgggca cagtggagag ccatgagtgg gggctgtcct gggggctgcg ctatggtgag 2040
tgagcccacg ccaggagcca gttcagccca gcctcattct tctcctgctt ctttgctttc 2100
caagttctag ttcataagca accagcacag tacctggaat gtaagaggcg cttggcttat 2160
aaaagagaga agaaaggaac ccctttcacc acgagggacg tcaggcccag ctctctgtcc 2220
cgggcattca tctttttccc ggcaattagc ttcagctcta atgacctgct gtctgcttta 2280
tgttgcacct atgatttcca gattgtaaag tttgtggcaa atgtattaac gttcacatcc 2340
catttacaga ttagcaaact aaggatctga gaggttatgt gattttccca gtctggcgga 2400
gggcagggcc ctgtcagaag ccgagacaca cctcgtaaca gcttgttatt tggggcttcc 2460
catcagtgag gcaggagtca taatggggcc ctggcttctg ggcaattccc ataaaacatt 2520
gtgccacctg aagatggaag tgtcagagga aggcgagatg ttatttctga aagcatctac 2580
cctcagaaag agacagagga gtgcatcgca gctgtgtgaa cccggggtgg tggtggttcc 2640
agcagcgggg cacgtggtgg acgcaagacg gcccgtttg agttttgctg ctaggtgaca 2700
cgtagaactt gaacttcaga gggtgctcct tggaggccct gatgagagag acacagaagg 2760
gaagggagac tgttcccatg gtcctgcctc tgcggcagcg agcctgctgc tcttgcagct 2820
gggtggctgt tcggtgcagg gggccggtga ggaacacggt ggggagcagc ctgcgagggc 2880
gccgggcttt cagcactcac tatacctgct gtttggtttg gccaggatgg cctgggagtc 2940
cccagggccg ggcaaagact gcccccaaca gacatgggcg gagggaaccg tctcacccct 3000
gcctcttctc ccttctctac gcagtgtcag tgggacggga gggtggtctg cccggggctt 3060
caagctcctg tccaggaacc ggacccatgt cgtacgccag tgcagccaca cggccagctt 3120
cgaggtgctc atggacgtct ccaggtgtga gatgggcatc ttcttgtggc cactatgccc 3180
gctgagccta cccccaaacc tgcccctcct caaactaaat cttccattct aagcactttg 3240
aacgaagcaa atcaaccaat taaaaaaata tattcccaga 3280

<210> 77

<211> 5567
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503990CB1

<400> 77
attcgcgcgc gccttcccta gccacccggg gttgcctcct aacatggaat ggccaaagga 60
gcgccccttg cgggaagtga gggtgggttg ggactgggtc cgcgttgggg gaggtgcaat 120
cttcgggttt cgcctctcgg ctccctctgg ctctggagtt gggacccctt gtgggctctg 180
gaagtccgcc tgagacttgg gtcaaggagt caaactgtcg ccccccgctc ctcccccaga 240
aatccggtga gcggtaagga aagtgatgcc aagtcttcga agcctcagtg acaaacgcat 300
agcaagaaca catccactcc agagatacct tctcgaaaca aaagattttc ctacctgctt 360
atacttggta accgagggaa ttactaagac ttcttgctca tttctgagta ttgtctttat 420
atcctgacac tatgaatgct acttggatgc ctcttaagtc tgttctctgg ggaggcagta 480
aggggccgtg gagctggcct cggcctcggc atcgggagag gctggacttc ctgtctctct 540
gtgctgaatg gctgcgatgg cgcccgctct cactgacgca gcagctgaag cacaccatat 600
ccggttcaaa ctggctcccc catcctctac cttgtcccct ggcagtgccg aaaataacgg 660
caacgccaac atccttattg ctgccaacgg aaccaaaaga aaagccattg ctgcagagga 720
tcccagccta gatttccgaa ataatcctac caaggaagac ttgggaaagc tgcaaccact 780
ggtggcatct tatctctgct ctgatgtaac atctgttccc tcaaaggagt ctttgaagtt 840
gcaaggggtc ttcagcaagc agacagtcct taaatctcat cctctcttat ctcagtccta 900
tgaactccga gctgagctgt tggggagaca gccagttttg gagttttcct tagaaaatct 960
tagaaccatg aatacgagtg gtcagacagc tctgccacaa gcacctgtaa atgggttggc 1020
taagaaattg actaaaagtt caacacattc tgatcatgac aattccactt ccctcaatgg 1080
gggaaaacgg gctctcactt catctgctct tcatgggggt gaaatgggag gatctgaatc 1140
tggggacttg aaggggggta tgaccaattg cactcttcca catagaagcc ttgatgtaga 1200
acacacaatt ttgtatagca ataatagcac tgcaaacaaa tcctctgtca attccatgga 1260
acagccggca cttcaaggaa gcagtagatt atcacctggt acagactcca gctctaactt 1320
ggggggtgtc aaattggagg gtaaaaagtc tcccctgtct tccattcttt tcagtgcttt 1380
agattctgac acaaggataa cagctttact gcggcgacag gctgacattg agagccgtgc 1440
ccgcagatta caaaagcgct tacaggttgt gcaagccaag caggttgaga ggcatataca 1500
acatcagctg ggtggatttt tggagaagac tttgagcaaa ctgccaaact tggaatcctt 1560
gagaccacgg agccagttga tgctgactcg aaaggctgaa gctgccttga gaaaagctgc 1620
cagtgagacc accacttcag agggacttag caactttctg aaaagcaatt caatttcaga 1680
agaattggag agatttacag ctagtggcat agccaacttg aggtgcagtg aacaggcatt 1740
tgattcagat gtcactgaca gtagttcagg aggggagtct gatattgaag aggaagaact 1800
gaccagagct gatcccgagc agcgtcatgt acccctgaga cgcaggtcag aatggaaatg 1860
ggctgcagac cgggcagcta ttgtcagccg ctggaactgg cttcaggctc atgtttctga 1920
cttggaatat cgaattcgtc agcaaacaga catttacaaa cagatacgtg ctaataagat 1980
tgagtctgtt tctcagccat tggaaaacca tggtgcccct attattggtc atatttcaga 2040
gtcactgtct accaaatcat gtggagcact cagacctgtc aatggagtta ttaacactct 2100
tcagcctgtc ttggcagacc acattccagg tgacagctct gatgctgagg aacaattaca 2160
taagaagcaa cgactgaatc tcgtctcttc atcatctgat ggcacctgtg tggcagcccg 2220
gacacgtcct gtactgagct gtaagaagcg gaggcttgtt cgacccaaca gcatcgttcc 2280
tctttccaag aaggttcacc ggaacagcac aatccgccct ggctgtgatg tgaatccctc 2340
ctgcgcactg tgtggttcag gcagcatcaa caccatgcct cccgaaattc actatgaagc 2400
ccctctgttg gaacgtcttt cccagttgga ctcttgtgtt catcctgttc tagcatttcc 2460
agatgatgtt cccacaagcc tgcatttcca gagcatgctg aaatctcagt ggcagaacaa 2520
gccttttgac aaaatcaaac ctcccaaaaa gttatcgctt aagcacagag cacccatgcc 2580
gggcagtctg ccagattcag ctcgtaagga caggcacaaa ttggtcagct ccttcctaac 2640
aacagccatg ttgaagcatc acacagacat gagcagttcg agctacttgg cagccaccca 2700
ccatcctcca cacagtccct tggtgcgaca gctctccacc tcctcagatt cccctgcacc 2760
cgccagctct agctcacagg ttacagccag cacatcgcag cagccagtaa ggaggagaag 2820

gggagagagc tcattcgata ttaacaacat tgtcatccca atgtctgttg ctgcaacaac 2880
tcgcgtagag aaactgcaat acaaggaaat ccttacgccc agctggcggg aggttgatct 2940
tcagtctctg aaggggagtc ctgatgagga gaatgaagag attgaggacc tatccgatgc 3000
agccttcgcc gccctgcatg ccaaatgtga ggagatggag agggcacggt ggctgtggac 3060
cacgagtgtg ccaccccagc ggcggggcag caggtcctac aggtcatcag acggccggac 3120
aacccccag ctgggcagtg ccaacccctc caccccccag cctgcctccc ctgatgtcag 3180
cagtagccac tctttgtcag aatactccca tggtcagtcc cctaggagcc ccattagccc 3240
ggaactgcac tcagcacccc tcacccctgt ggctcgggac actctgcgac acttagccag 3300
tgaggatacc cgttgttcca caccagagct ggggctggat gaacagtctg tccagccctg 3360
ggagcggcgg accttccccc tggcgcacag tccccaggcg gagtgtgagg accagctgga 3420
tgcacaggag cgagcagccc gctgcactcg acgcacctca ggcagcaaga ctggccggga 3480
gacagaggca gcgcccacct cgcctcccat tgtccccctc aagagtcggc atctggtggc 3540
agcagccaca gctcagcgcc cgactcacag atgagcggga gacagccatc taaacagact 3600
cactaactat tggcattaaa gcttcagaaa tctctgcgtt tgatattcaa acatcatatg 3660
ccggaaattt tcacagtttt tagtgaactt aaggaattta gatcctactt tggtattttt 3720
ttttcttgtt ttaatttttg ttttgttttt gtttccatgt tttcttgtca cacacctgag 3780
cacttcctcc cgttggcaaa cagaagttca ggatgagacc ctgctggcct ggtcctgggc 3840
acatcctctg cactgttgaa tcactggact tactgatctt agatgaccac cccctccctc 3900
acacctgtgg gcagggcaga acagcctggc gggctacagt ttagcatggc cttcttgagc 3960
tagggtggaa tggggcaggg tgctctggac tcttacccc tcccctccca tctgtggctt 4020
ggctctgctg tggccctcct ggctgggtcc ccttggtttt tcgtgctgga acatccccac 4080
cagagcctct ctgccataac tgccagctgc tctccccgag tgctcagctg gcagaacacc 4140
tttcctttct cacccagaac ttaagagact gattttttgt ttcatctgca tttggtcttc 4200
tctgttttga ctctttcact gcagtaacct ggctgtggct gctcaggttc ccctcctcat 4260
gccccttggt acccttccct gtctgctctc ccatgccatg tacacaccca caacccgtcc 4320
ttccacttgg aatattttta ccacctatcc tgatctttga aggtagggtt aggactactt 4380
aacctctatt cccactcccc tgcaaactgg gggttgtggg aagtgagcag ccatctccct 4440
gtgtgatttt tttttttttc cctctgattc actttgccat gtttccttca catccagatc 4500
cctgtcggtg ttagttccac tcttggtctt tcacgctccc cttgcctgtg gaacattgtc 4560
tggtcctagc tgtggttccc attgttcccc cttcaccctt ctctgttaac cttgtgcctg 4620
tctcctgtat gatcacatca ccaaaaaggg ggaggggga gaagactctt ttttttggc 4680
cattttgtaa tcgtataaaa atagtagaca actgcttaat ggttggggtt ttttcacaat 4740
tttcaacatt agtgattttt ttttctgttt gcaagttaaa gggtttgtca ttgtttcttt 4800
aaaaaaaaaa tacaataatg caccatatcc ctatgcataa agtgcttctt ctatttataa 4860
ggttgaaaat tctgaataac ccttttagca ttgaaaaaaa aaacaaaaac aaaaaatgga 4920
aaaaaaaaac cttgtatttt gtaaatattt tcttttcctg ctttggagct gtgtaatggc 4980
agcgaaacat gtagctgtct ttgttctata gaaatgcttt tcttcagaga agctgatctt 5040
tgttaatgtc ttgattctgt tcgcaaagca cagactagtg cttaaaaaaa aaaaagaagg 5100
aaaaattgaa aaaaataaaa aaaaaagtta cagaaaaaaa aaaaaggggg gggggtaccc 5160
gccccgaatt cgaaaccatg taaaagctgt tcccggggga aatttttac ccgcccaaat 5220
ttccacacaa cttaggaccc ggaaccgcaa gggttaaacc ctggggtccc aaatgggtga 5280
cctaccccg ttaattggct ttgcgctcac tgcccgtttt ccagtgggga accctgtgtt 5340
cccagcgcct ttaaggatcc gcccaacgcg gggggatacg cgttttcctt ttggggcgcc 5400
tttccgtttc cgggcccact aacgcgctgc cccgggtctt tcggcttcgg cggggggtat 5460
cagctcactc aaagggggta aaagggtttt ccaaagattc gggggttaac cctggaaaga 5520
acttttggcc caaagcccag ctaaaggccg ggacccttaa aaaggcc 5567

<210> 78
<211> 1263
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504655CB1

<400> 78
ttcagggcgt atcggtaaca gcaccctacc ctttgcaaaa atctccttgg gaaccagaaa 60
cttaaacaca accaggaaga aaaaaaatca gccaaaaata aaagcgaatt aagacagttg 120
gggtcttatt ttagaaatat accttctagg ttctggtatg ttgggctctg tcatctttta 180
acccggatca aggaaccact gtgacatttt ttggtaactt ccagtctgta gtttggattt 240
aggagttcag gattccaatc ctaatttcca ggttccagct gcgtggcccg agtagcccgg 300
ccagccccac cccgggcgc aggaggcgcg cgaagggcgg gagttctcgg gtccccgccc 360
tcggggaggc tctgggcgcg gggcgggact ccgggacgct gggcgggcgt gctagggccc 420
cgccgccctg gcccggcctt gccttgcgct gcgcgctcac catggtgggc cccgcgccgc 480
ggcggcggct gcggccgctg gcagcgctgg ccctggtcct ggcgctggcc ccggggctgc 540
ccacagcccg ggccgggcag acaccgcgcc ctgccgagcg ggggccccca gtgcggcttt 600
tcaccgagga ggagctggcc cgctatggcg gggaggagag ttttatggac gaggagcccc 660
ctacaatgcc ttgacgggga aggactccac tagaggggta gccaagatgt ccttggatcc 720
tgcagacctc acccatgaca ctacgggtct cacggccaag gaactggagg ccctggatga 780
ggtcttcacc aaagtgtaca aagccaaata ccccatcgtc ggctacactg cccggagaat 840
tctcaatgag gatggcagcc ctaacctgga cttcaagcct gaagaccagc cccattttga 900
catcaaggat gagttctgat gttccccctg caggagcagg ttcttgggag cgtgaggcag 960
gaagacacta ggtgctgaat ctcctgcaaa actggctgcc tggaggccct gagccaccca 1020
gatctgaata aaacagatgc ttaccctgga aaaaaaaaaa aaaaaaaagc cccaaacttt 1080
agtgaagagc ccctgaaaag atccttgagg accactttgg gcctattcag aagaggtaca 1140
cagccatgcg atgactcctg ttcgccccca ggtggggaac ttgctctgta tctgaatagc 1200
tctttttgga ccccgtgaac atgcgggaat tcagaagact cgagatgcta tggtccctgt 1260
cca 1263

<210> 79
<211> 1854
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504690CB1

<400> 79
cggctcgagg ctgagaccca gagtcaccca ggggtctccg tcacgtgcca ggagtaggca 60
gaagtgggct gtgacagatc aggaaacaga gctcagtgca gcccactaaa ttgctcaggg 120
ccctacagct aacaagcggc agaggcagga tctgcactca ggagctgctt ggagatgctg 180
ctgtggccac tgctgctgct gctgctgctg ctgccaacat tggccctgct caggcagcag 240
cggtcccagg atgccaggct gtcctggctt gctggcctcc agcaccgagt ggcatggggg 300
gccctggtct gggcagccac ctggcagcgg cggaggctgg agcagagcac gctccatgtg 360
caccagagcc agcagcaggc cctgaggtgg tgtctacagg gagcccagcg cccccactgt 420
tccctcagaa ggagcacaga cataagcacc ttccggaatc atctccctct gaccaaggcc 480
agccagaccc agcaggaaga cagtgggcgg gggccaagct gctggaccat ggctgtgtgg 540
agagcagcat tctggattcc tctgcgggct ctgctcccca ctacgaggtg tttgtggcgc 600
tgaggggggct gaggaatctg tcagaggaaa atcgagacaa gctggaccac tgccttcagg 660
aagcctctcc ccgctacaag tccctgcggt tctggggcag cgtgggccct gccagagtcc 720
acctggtggg gcagggagcc ttccgagcac tccgggcagc cctcgctgcc tgcccctcct 780
ccccccttccc ccctgcgatg ccccgggtcc ttcggcacag gcacctggcc cagtgtctgc 840
aggagagggt ggtgtcctga gtcaagtcct gccccaccgc ccagctcccc ccagaggcca 900
cctcgcccct ccctctggga cctctccgga tggggagtcc ttggccaggg tctctgactc 960
tgtgtcacct gacatttgcc catgagagcc gctgggcctt agagaggcct tggcccagct 1020
gaccggttct gaagtatggg cctccggggt tagcagatgc cagcagtgcc tgcccgtgtc 1080
cccatgtccc ggcatgaagg acactgctag agagttacca tgcacaccga tggtttcctg 1140
tatcacagcc caaagaggtt ctctggtggc cacagctgtg tgctcagtca gtgcactggg 1200
caagctagaa gtgttggggg gttaatgtcc ccaggagcag caaccctgag tcaataagga 1260
gcaggacctc agcttcattg tccttgagca ggacaattct gaagtgtatt ctacataaac 1320

tctcagagga tgcccagcag gatggagtcc cagttgcccg cagcagtaac ccactcattc 1380
atgtacttcc tgcgggggct ctcccttccc tctcttcccc actcccccgc cttgggcttc 1440
ctgggatggc tcccaaataa acctcttgca cccagaaaaa aaaaaaaaaa aaaaaaaaaa 1500
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aggggggcc 1560
gctctagggg ttccaggttt aggtacgggt gcatgggagg tcatagctct tctaaggtgt 1620
cccctaattt catttcacgg gcggtggttt taaaaggtcg tgactgggaa aaccctgggg 1680
ttacccaatt taatcgcttt gaggaaattc ccctttggc aagttggggt aatagcgaag 1740
gggcccgcac ggttggcctt tccaaaaatt tggccctct gattggcgat tgggacgcgc 1800
ctgtaggggg gttttaagcc ggcggttttg gtggttacgc cggtgacccg ttac 1854

<210> 80
<211> 1332
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504720CB1

<400> 80
cggctcgagg cctggtgacc agaagtttgg agtccgctga cgtcgccgcc cagatggcct 60
ccaggctgac cctgctgacc ctcctgctgc tgctgctggc tggggggctg gggagaacac 120
caaaacaaac ctggagagca tcctctctta ccccaaggac ttcacctgtg tccaccaggc 180
cctgaagggc ttcacgacca aaggtgtcac ctcagtctct cagatcttcc acagcccaga 240
cctggccata agggacacct ttgtgaatgc ctctcggacc ctgtacagca gcagccccag 300
agtcctaagc aacaacagtg acgccaactt ggagctcatc aacacctggg tggccaagaa 360
caccaacaac aagatcagcc ggctgctaga cagtctgccc tccgataccc gccttgtcct 420
cctcaatgct atctacctga gtgccaagtg gaagacaaca tttgatccca agaaaaccag 480
aatggaaccc tttcacttca aaaactcagt tataaaagtg cccatgatga atagcaagaa 540
gtaccctgtg gcccatttca ttgaccaaac tttgaaagcc aaggtggggc agctgcagct 600
ctcccacaat ctgagtttgg tgatcctggt accccagaac ctgaaacatc gtcttgaaga 660
catggaacag gctctcagcc cttctgtttt caaggccatc atggagaaac tggagatgtc 720
caagttccag cccactctcc taacactacc ccgcatcaaa gtgacgacca gccaggatat 780
gctctcaatc atggagaaat tggaattctt cgatttttct tatgacctta acctgtgtgg 840
gctgacagag gacccagatc ttcaggtttc tgcgatgcag caccagacag tgctggaact 900
gacagagact ggggtggagg cggctgcagc ctccgccatc tctgtggccc gcaccctgct 960
ggtctttgaa gtgcagcagc ccttcctctt cgtgctctgg gaccagcagc acaagttccc 1020
tgtcttcatg ggcgagtat atgaccccag ggcctgagac ctgcaggatc aggttagggc 1080
gagcgctacc tctccagcct cagctctcag ttgcagccct gctgctgcct gcctggactt 1140
ggcccctgcc acctcctgcc tcaggtgtcc gctatccacc aaaagggctc cctgagggtc 1200
tgggcaaggg acctgcttct attagccctt ctccatggcc ctgccatgct ctccaaacca 1260
ctttttgcag ctttctctag ttcaagttca ccagactcta taaataaaac ctgacagacc 1320
atgaaaaaaa aa 1332

<210> 81
<211> 1435
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504722CB1

<220>
<221> unsure
<222> (1) ... (1435)

<223> a, t, c, g, or other

<400> 81
tgcctgcctg agtcacgtgt caggggggaag ctggaaggcg tcgttctcct ttcccagctc 60
tcctgcctgt ccgccatgtt ttcaggccgg gtctggcttg gtcttccccc gtaaggaaat 120
ggccggggag ctccagggga cccaggcgcc gtcgcttcgg cggagcctgg gctgaccagc 180
caggacagcg gggtaaaccc gaacaattct gcgcgaggta gggaggccat ggcgtccggc 240
agtaactggc tctccggggt gaatgtcgtg ctggtgatgg cctacgggag cctggtgttt 300
gtactgctat ttatttttgt gaagaggcaa atcatgcgct ttgcaatgaa atctcgaagg 360
ggacctcatg tccctgtggg acacaatgcc cccaaggttg ctacaactat ctgtatagga 420
tgaaagctct ggatgccatt cgtacctctg agatcccatt tcattctgaa ggccggcatc 480
cccgttcctt aatgggcaag aatttccgct cctacctgct ggatctgcga aacactagta 540
cgcctttcaa gggtgtacgc aaagcactca ttgataccct tttggatggc tatgaaacag 600
cccgctatgg gacaggggtc tttggccaga atgagtacct acgctatcag gaggccctga 660
gtgagctggc cactgcggtt aaagcacgaa ttgggagctc tcagcgacat caccagtcag 720
cagccaaaga cctaactcag tcccctgagg tctccccaac aaccatccag gtgacatacc 780
tcccctccag tcagaagagt aaacgtgcca agcacttcct tgaattgaag agctttaagg 840
ataactataa cacattggag agtactctgt gacggagctg aaggactctt gccgtagatt 900
aagccagtca gttgcaatgt gcaagacagg ctgcttgccg ggccgccctc ggaacatctg 960
gcccagcagg cccagactgt atccatccaa gttcccgttg tatccagagt tcttagagct 1020
tgtgtctaaa gggtaattcc ccaacccttc cttatgagca tttttagaac attggctaag 1080
actattttcc cccagtagcg ctttttttctg gatttgcatt caggtgttat tcttaatgtt 1140
tctgtcaaag cttcttaaaa atcttcactt ggtttcagcc atagttcacc ttccctgttc 1200
caggtttatt taattccaaa ggtgagagtt ggagtgagat gtcttccata tctatacctt 1260
tgtgcacagt tgaatgggaa ctgtttgggt ttagggcatc ttagagttga ttgatggaaa 1320
aagcagacag gaactggtgg gaggtcaagt ggggaagttg gtgaatgtgg aataacttac 1380
ctttgtgctc cactttccag cacactgcgc cgnttaactc gtgagtccgg ctcgg 1435

<210> 82
<211> 554
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504733CB1

<400> 82
ggaggcaaga tgtgtgagca gcatccaaag aggcctgggc ttcagttgtg gagagggaga 60
gagccaggtt ggaatgggca gcaggtaggg agatccctgg ggaggagctg aagcccattt 120
ggcttcagtg tcccccaaac ccccaccacc ctcttcctag gccgccctcc ccactgttac 180
caacatgaag ctgctcgcag caactgtgct actcctcacc atctgcagcc ttgaagtact 240
tccagaccgt gactgactat ggcaaggacc tgatggagaa ggtcaagagc ccagagcttc 300
aggccgaggc caagtcttac tttgaaaagt caaaggagca gctgacaccc ctgatcaaga 360
aggctggaac ggaactggtt aacttcttga gctatttcgt ggaacttgga acacagcctg 420
ccacccagtg aagtgtccag accattgtct tccaacccca gctggcctct agaacaccca 480
ctggccagtc ctagagctcc tgtccctacc cactctttgc tacaataaat gctgaatgaa 540
tccaaaaaaa aaaa 554

<210> 83
<211> 2365
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7507100CB1

<400> 83
agactctgcg cacgcgcagt tccccctccc gcccctgctt gccggtgatg gcgcataacg 60
catgcgcggg gagggcggag ctgggcgttg ccgtggctac tgggaacgca tttcacgggg 120
gcggggcgtg gttccggggc ggggcgcggc cgccggaagt gcgtggccgc ccggggccat 180
ggcgacactc agcttcgtct tcctgctgct gggggcagtg tcctggcctc cggcttctgc 240
ctccggccag gagttctggc ccggacaatc ggcggccgat attctgtcgg gggcggcttc 300
ccgcagacgg tatcttctgt atgacgtcaa ccccccggaa ggcttcaacc tgcgcaggga 360
tgtctatatc cgaatcgcct ctctcctgaa gactctgctg aagacggagg agtgggtgct 420
tgtcctgcct ccatggggcc gcctctatca ctggcagagt cctgacatcc accaggtccg 480
gattccctgg tctgagtttt ttgatcttcc aagtctcaat aaaaacatcc ccgtcatcga 540
gtatgagcag ttcatcgcag aatctggtgg gccctttatt gaccaggttt acgtcctgca 600
aagttacgca gaggggtgga aagaagggac ctgggaagag aaggtggacg agcggccgtg 660
tattgatcag ctcctgtact cccaggacaa gcacgagtac tacaggtgcc tgctgaggct 720
gctccccctg ccgcagggct cagcctccat cgtggcgccc ctgctgctga gaaacacatc 780
agcccggtcc gtgatgttag acagagccga gaacctactt cacgaccact atggagggaa 840
agaatactgg gatacccgtc gcagcatggt gtttgccagg cacctgcggg aggtgggaga 900
cgagttcagg agcagacatc tcaactccac ggacgacgca gacaggatcc ccttccagga 960
ggactggatg aagatgaagg tcaagctggg ctccgcgcta gggggcccct acctggagt 1020
ccacctgaga agaaaagatt tcatctgggg tcacagacag gatgtaccca gtctggaagg 1080
ggccggagga acgggcaggc cgctggcctg aggtctggag aaacccctgg agaagggtgt 1140
ccccaccagc ccatacagcg tgtgtgtgga ggggccttg acctccgtga tgtctactgt 1200
gcctcaggat aaggacccgc catgccctgg ctagacagtg tgtggttttt tattggcacc 1260
tcagtctcaa catttctttt tcggattcat gaggaaagag aaatcctggg gttggacccc 1320
aagacgacgt acaacaggtt ctgcggagac caagagaagg cgtgtgagca acccacccac 1380
tggaagatca cctactgagg aggatcctcc agggccgctc cccggacccg acaggcgcgg 1440
gtggatgcag gttctgtcgc cgtggagtca ccgtctactg ccagccggga gctgggcgga 1500
caggaccgtc cctcgcaggg tcccaggccc agaagaggcc ccacgcctct agagctgggc 1560
tccgtcctcg gcgttgccag ccgccatggc tgatgaagag gctccgctgc tctcgggggt 1620
ggcggttgtt ttcaggcagc gtctgtgaac ccacagctcg gttgccagca gtgcccgcgt 1680
ggtgacccag aagcaggagt gtttgtcagg ctcccgctct ggcctttcca gccacctttc 1740
atgtcttcat attttaagtg cattgaggat agatgcaggc gggtgagctg ccctccgtca 1800
ggtggacccg ggctgacatt tccctgggag ctggtgcaag gagaagcgtc attttaaatg 1860
tctgcagagc gaccaggggc ctcatgaatc tctccgttgc cctccgcgca gcaggaggct 1920
gcctgtgtgt ttcctcctgg gatgcgtgca aggcagacct ggtgctgcaa aggaaagggc 1980
ctgaggcctc agggagcccc gtggagggat gacagttcag gccctactgc tggcacgtca 2040
gagcactggg aagtttttca gtgacgtctc tggggcactc agtggattgt ctgtaggaaa 2100
cttgcagctc tgctcctcac accaggcccg gctggccacc caccctcgcc cccactggcc 2160
acccctccct cgccccgact gccccgcccc accctcaccc cgactgcccc gccctcgccc 2220
ggctggccgt ccctgccctc gccccggctg gcaggtgcac atggggcctc caggtctgcc 2280
attcgctatt gagaactaga aatgaggaag gacagttacg ctaactccaa aaggctgtct 2340
aggatgagct gctttatcag ggagc 2365

<210> 84
<211> 703
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503330CB1

<400> 84
gccgaaggcc agctggagcg tcggcgctgc ggggccgggg gggtcgaatg ttcgtggcat 60
cagagagaaa gatgagagct caccaggtgc tcaccttcct cctgctcttc gtgatcacct 120

cggtggcctc tgaacacgcc caactacttc gacacgtcgc agcccaggat gcgggagacg 180
gccttcgagg aggacgtgca gctgccgcgg gcctatatgg agaacaaggc cttctccatg 240
gatgaacaca atgcagctct ccgaacagca ggatttccca acggcagctt gggaaaaaga 300
cccagtggca gcttggggaa aagacccagc gctccgttta gaagcaacgt gtatcagcca 360
actgagatgg ccgtcgtgct caacggtggg accatcccaa ctgctccgcc aagtcacaca 420
ggaagacacc tttggtgaaa gactttaagt tccagagaat cagaatttct cttaccgatt 480
tgcctccctg gctgtgtctt tcttgaggga gaaatcggta acagttgccg aaccaggccg 540
cctcacagcc aggaaatttg gaaatcctag ccaaggggat ttcgtgtaaa tgtgaacact 600
gacgaactga aaagctaaca ccgactgccc gcccctcccc tgccacacac acagacacgt 660
aataccagac caacctcaat ccccgcaaac taaagcaaag cta 703

<210> 85
<211> 3871
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504519CB1

<400> 85
cccacgcgtc cggcagctga gacggcagcg gcagcttctc agggccggag ccagttcttg 60
gaggagactc tgcacagggc atggatcact gtggtgccct tttcctgtgc ctgtgccttc 120
tgactttgca gaatgcaaca acagagacat gggaagaact cctgagctac atggagaata 180
tgcaggtgtc caggggccgg agctcagttt tttcctctcg gcaggaggtc agcatgcccg 240
gggtcagcac gccatgcagt tccccgccga gctgacccgg gacgcctgca agacccgccc 300
cagggagctg cggctcatct gtatctactt ctccaacacc cactttttca aggatgaaaa 360
caactcatct ctgctgaata actacgtcct gggggcccag ctgagtcatg ggcacgtgaa 420
caacctcagg gatcctgtga acatcagctt ctggcacaac caaagcctgg aaggctacac 480
cctgacctgt gtcttctgga aggagggagc caggaaacag ccctgggggg gctggagccc 540
tgagggctgt cgtacagagc agccctccca ctctcaggtg ctctgccgct gcaaccacct 600
cacctacttt gctgttctca tgcaactctc cccagccctg gtccctgcag agttgctggc 660
acctcttacg tacatctccc tcgtgggctg cagcatctcc atcgtggcct cgctgatcac 720
agtcctgctg cacttccatt tcaggaagca gagtgactcc ttaacacgca tccacatgaa 780
cctgcatgcc tccgtgctgc tcctgaacat cgccttcctg ctgagccccg cattcgcaat 840
gtctcctgtg cccgggtcag catgcacggc tctggccgct gccctgcact acgcgctgct 900
cagctgcctc acctggatgg ccatcgaggg cttcaacctc tacctcctcc tcgggcgtgt 960
ctacaacatc tacatccgca gatatgtgtt caagcttggt gtgctaggct gggggcccc 1020
agccctcctg gtgctgcttt ccctctctgt caagagctcg gtatacggac cctgcacaat 1080
ccccgtcttc gacagctggg agaatggcac aggcttccag aacatgtcca tatgctgggt 1140
gcggagcccc gtggtgcaca gtgtcctggt catgggctac ggcggcctca cgtccctctt 1200
caacctggtg gtgctggcct gggcgctgtg gaccctgcgc aggctgcggg agcgggcgga 1260
tgcaccaagt gtcagggcct gccatgacac tgtcactgtg ctgggcctca ccgtgctgct 1320
gggaaccacc tgggccttgg ccttcttttc ttttggcgtc ttcctgctgc cccagctgtt 1380
cctcttcacc atcttaaact cgctctacgg tttcttcctt ttcctgtggt tctgctccca 1440
gcggtgccgc tcagaagcag aggccaaggc acagatagag gccttcagct cctcccaaac 1500
aacacagtag tccgggcctc ctggcctgga atcctcagcc tctctggccg ccagtagcct 1560
gaggctacgg ctcctgctag agagggtggc aggcctgctg ctggacccca gaggccactg 1620
tgaccgccaa ggggcctttt ccacttccac ggcctctcca ggcactgagg ggaaggcatt 1680
gctctacctc tccctgacat tttgctccgg ggcagatcca accttacctg gggcagcaaa 1740
ctttgtcctg gtacctgggc ccagctcgcc agggatgtgg gcagagcacc agcctgggca 1800
tcaggaagcc aagtttcaag gactgtcttt gagtctgtct gtatgacctt gggcctgcca 1860
cttctcacag accctaggta tccacagctg tgacatgggg gcaagcagct ttgtttcagc 1920
ctaacccagg agcttagtaa aaattgcata agaccagggg gaagagtgtc agcgtggggt 1980
gggaattccc gcggcctcca cctgcttgct aggggcagga tctcattcag gctgccctgg 2040
aagcacctgc ttggccctgc caccttcctc caggggaggg ccagatggca tcctggcttg 2100

```
gggcgggtgg gacctacccc aggctctgag actttactgg cctatgcctg aggcctcttt 2160
tcctttaact ccctaaatta tgatgactcc aagtccaagc ccacccttcc caaagattgg 2220
gaggttccgc cgttcccaga ggctcctcct gcggtgctcc caagacttcc atagaccatc 2280
tggaccagta gcccatcccg cagttttctt gggggcagag gaaaacgctt ctttctcctc 2340
cagctgaatc agctggatcc cagtgtcctg gctgtttggt gattgggcaa gattgaattt 2400
gcccaggtag gcgtgagagt gtgggtttta aattcgaagc tcaggccata gtttcagaga 2460
atcaccctta ccccagacct tcatgagaca gtgctcatga agccagtgcg tttcccagaa 2520
cgaacactag gcggcaccgt tggtccacac tcagaggccc ttggcgccaa gactgcatct 2580
agaatcgctc aaacacctgt ttgcagaccc catgcaccag ctggaggggc cgtaactgca 2640
ggactgcgcc tactgagtga cccatttcct ccaggaggaa aggcaagaca cgcttacacg 2700
gccatttgtc tcttttccca atgcggcggt gcactttcgc tcttgggggc tgcaccccag 2760
acatagctgg caccagagca gggtgctcag gtggtgggtg ctcagggccc tgccccaggc 2820
cactgggccg ttttgatgac ctcgaaggtc acaggcagaa aataggagca ggatttcccc 2880
tggggaaaag ttctcctggg acatcttctg ctcttctgta catttctaga tgcaaataac 2940
tccttcacca ggcagtgagt ggcgtaggct ctggagccag gctgcctggg ctccaatgcc 3000
agctctgcca cttgctagct gtgagactgt ggacaaacca ctcagcctct gtgtgcctca 3060
gttttcctat ttgtaaaata gaggccatag tggtacctat tttgaagact aagtaaaaga 3120
attcaaataa agagacttgg cacagagtaa gtgctcagta aaaaaaaaaa aaaaaaaaaa 3180
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaac acaagagggg 3240
ggcgggggcg cgagcgcgaa aatagggcag ggtgcagccc gcgggcggga ataaatacta 3300
cggcgcgcgc gttagctagt gggaggcgca gagaagaaaa ggaggggacg caccgcccga 3360
tagagaaaaa gtaagaggaa agaaacaact accaggaggt agggcaaccg ggggaggggt 3420
ggaaccagga cagccaggca gaaagaacac tctgggagga gaggcaacaa cacaataata 3480
gacagccgct gtgaagaaaa cgcggcgggt agtataccca ccccaaatgt ggtgtgtcaa 3540
cccccgcggg gagagcgccc accggtgggg gacaacactc ggggaaagaa acaactcgcc 3600
gtgagaaaag ggtggtgaga ggatccaaaa agcggtacac agtgtgcatg aacgcacacg 3660
aagatagggg agaaggaaca accccgcgcc ggacaggaaa agcggggcgc ggcacacagg 3720
agcgggatca gcgcccaagt gaaaagtgac gagggcgcgc caaaagtggt aagcgcacca 3780
accgcggca cgttggggtc gggaacaacg cgtacgacga agggcaacag tgcacaccgc 3840
gatcgctgag ttgaggacag acgacactga t                                 3871
```

<210> 86
<211> 695
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504705CB1

<400> 86

```
ccgggactcc aacatgcaat gattgcccca gatctctgta ctacacctga ccactttctg 60
aaatatttga agccgaggat acaggaagtt cttaaaggag cctctctccc tactgctgct 120
acacaagacc ctgagactga cctgcaggac gaaaccatga agagcctgat ccttcttgcc 180
atcctggccg ccttagcggt agtaactttg tgttatgaat cacatgaaag catggaatct 240
tatgaactta gatccgagaa cgctctaagc ctgtccacga gctcaatagg gaagcctgtg 300
atgactacag actttgcgaa cgctacgcca tggtttatgg atacaatgct gcctataatc 360
gctacttcag gaagcgccga gggaccaaat gagactgagg gaagaaaaaa aatctctttt 420
tttctggagg ctggcacctg attttgtatc cccctgtagc agcattactg aaatacatag 480
gcttatatac aatgcttctt tcctgtatat tctcttgtct ggctgcaccc ctttttcccg 540
cccccagatt gataagtaat gaaagtgcac tgcagtgagg gtcaaaggag agtcaacata 600
tgtgattgtt ccataataaa cttctggtgt gataaaaaaa aaaaaaaaaa aaaaaaaaaa 660
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa                             695
```

<210> 87
<211> 536

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504738CB1

<400> 87
gctaggttct gaagcttctg agttctgcag cctcacctct gagaaaacct ctttgccacc 60
aataccatga agctctgcgt gactgtcctg tctctcctca tgctagtagc tgccttctgc 120
tctccagcgc tctcagcacc aaattccaaa ccaaaagaag caagcaagtc tgtgctgatc 180
ccagtgaatc ctgggtccag gagtacgtgt atgacctgga actgaactga gctgctcaga 240
gacaggaagt cttcagggaa ggtcacctga gcccggatgc ttctccatga gacacatctc 300
ctccatactc aggactcctc tccgcagttc ctgtcccttc tcttaattta atctttttta 360
tgtgccgtgt tattgtatta ggtgtcattt ccattattta tattagttta gccaaaggat 420
aagtgtcccc tatggggatg gtccactgtc actgtttctc tgctgttgca aatacatgga 480
taacacattt gattctgtgt gttttcataa taaaacttta aaataaaatg caaaaa 536

<210> 88
<211> 549
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510280CB1

<400> 88
agccgagctg ggcgagaagt aggggagggc ggtgctccgc cgcggtggcg gttgctatcg 60
cttcgcagaa cctactcagg cagccagctg agaagagttg agggaaagtg ctgctgctgg 120
gtctgcagac gcgatggata acgtgcagcc gaaaataaaa catcgcccct tctgcttcag 180
tgtgaaaggc cacgtgaaga tgctgcggct gcattgcttg gtagtcttca tgtaaggctc 240
agctttattg tgtgttttta ggatattatc aactcactgg taacaacagt attcatgctc 300
atcgtatctg tgttggcact gataccagaa accacaacat tgacagttgg tggaggggtg 360
tttgcacttg tgacagcagt atgctgtctt gccgacgggg cccttattta ccggaagctt 420
ctgttcaatc ccagcggtcc ttaccagaaa aagcctgtgc atgaaaaaaa agaagttttg 480
taattttata ttacttttta gtttgatact aagtattaaa catatttctg tattcttcca 540
aaaaaaaaa 549

<210> 89
<211> 1888
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503700CB1

<400> 89
gcctgtgcgc tgtgcctccg cttgtctgct ctccccccgcc ttggccttct ctcgccgccc 60
cctttcgcgt cctctgtcct tgtgtcctct ctctcccgtc tcgctctcgt gtccttggac 120
tgctccttct gtctcgcgcg ctccttggtc gctccctccg tgttgctctg ttgtgtggtt 180
agctgtcttc cttttctctt gttccgtcgc tctttgcttc ttttcctttg gttcgttgtg 240
tctgtgcctg tctctggctt tcttggtttt tgtgcccagg cccccacgtt ggccgctgct 300
gccgggctca ccccagcccc gcccggaggc gcccccgcggc cccggctagc cagggcgggc 360
ggccacactc tgccctactc cttccctccg cgttccgggc ctcggagccg ccttgaggag 420

gatgagtccc tggagctggt tcctgctgca gaccctctgc ctcctgccca cgggcgcagc 480
ttcgcggcgc ggggcgcccg gcaccgccaa ctgcgagctc aagccccaac aaagcgagct 540
gaattccttc ttgtggacca ttaagcgaga cccaccatct tacttctttg gcacaatcca 600
tgtcccgtac acccgagttt gggacttcat ccccgacaac tctaaggagg ctttcctgca 660
gagcagcatt gtgtactttg agttggatct cacagacccc tataccatct cagctctcac 720
cagctgtcag atgctgccac agggcgagaa cctccaagat gtgctcccca gggacatcta 780
ctgccgcctc aagcgccacc tggagtatgt caagctcatg atgcccttgt ggatgacccc 840
agaccagcgc ggcaagggc tctacgcaga ctacctcttc aatgctattg ccggaaactg 900
ggagcgcaag aggcctgtct gggtgatgct catggtcaac tccctgactg aagtggacat 960
taagtcccgt ggagtgcctg tcttagacct gttccttgcc caggaggctg agcggctgag 1020
gaaacagact ggggcagtgg aaaaggtgga agagcagtgc catccattga atgggttgaa 1080
cttttcacag gtcatttcat gggcaacaac acagtgctgg atgttttgcg gcgtgaaggc 1140
tatgaggtag aacacgcccc tgctggacga cccatccaca aagggaagag taaaaagacc 1200
tccacacggc ccactctgtc caccatcttt gctccaaaag tccctaccct ggaagtaccg 1260
gcaccagaag ccgtatcctc agggcactca acgctgcctc cccttgtgtc ccggcctgga 1320
agtgccgaca cgcccagtga ggccgaacag aggttccgga agaagcggag gcggtcacag 1380
cggaggccgc gactccggca attcagcgat ctgtgggtcc gcctggagga gagtgacata 1440
gtcccgcaac tccaggtccc tgtcctggac aggcacatct ccactgaact gcggctccct 1500
cgccgtgggc attcccacca cagccagatg gtggccagca gtgcctgcct gtctctctgg 1560
actcctgtgt tctgggtgct ggtgctggct ttccaaacag agacacccct cctgtaacga 1620
ctggaagcac caggctaaga acctgacccc tcggacttga agaatggcca ttcctgtact 1680
ccacattctg gtctagcctt gttgggccca atccagaaga gactgctctt gaaaaagcgg 1740
cccagtgttg attcttctct ttccaaggaa tgtgactttg ggttatccaa ctttgggggc 1800
aggtgtacag ttttgtaaca tagtgagttg tgtgaaaata aattataaat gagttgtatg 1860
aaaataaata cctttttttg tataaaaa 1888

<210> 90
<211> 1050
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504685CB1

<400> 90
tacggggcgc ctgtgcttga ggttgggggt tgcgtcgctc tctggtaaag gcgtgcaggt 60
gttggccgcg gcctctgagc tgggatgagc cgtgctcccg gtggaagcaa gggagcccag 120
ccggagccat ggccagtaca gtggtagcag ttggactgac cattgctgct gcaggatttg 180
caggccgtta cgttttgcaa gccatgaagc atatggagcc tcaagtaaaa caagtttttc 240
aaagcctacc aaaatctccc tactgccaat aaagggaaaa taagagatgc tcatcgacga 300
attatgcttt taaatcatcc tgacaaagga ggatctcctt atatagcagc caaaatcaat 360
gaagctaaag atttactaga aggtcaagct aaaaaatgaa gtaaatgtat gatgaatttt 420
aagttcgtat tagtttatgt atatgagtac taagttttta taataaaatg cctcagagct 480
acaattttaa aaaatgattt agcacaagct aaatctcaaa gccttggtat aattttcttg 540
tttaaatttg ggattttaa atcagattat agtttagaat atttgcatat taattatggg 600
caagcacaca ccttctgaat agaaatattg ttcattactc atttagcaga taatttggga 660
cctatgtcta cttttcaagg caaagtgaag atgacagtcc ttgctctcag ggagccccca 720
ctttaatggg agactgataa actggtaatt agactgtgat aaatagtatg atggaaatta 780
gcttaagctg tttaagtagg gactcttctt attcggtgga aaggctgttc caggtacagg 840
caactggcct ggcaacttgg atacttggaa ccttgtattt aaaagtgaat ttaaccacaa 900
ctgagaccta agaaattgac ctaggggtgt gtgtgtgtgt gtattctatg tacatataaa 960
ccatttttat ttcatgcatt aaaaatagta tgataaagat ttcagagtac aggtctggta 1020
aatcacagtt cattgcagcc tcaacctcct 1050

<210> 91

<211> 899
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506844CB1

<400> 91
cgtgaactgg gcagctgtga cacgcgcacc ttgtcacctt gtcaccaatt tcctgcccgg 60
gatcccagag ctacccttct tggagcccac aggggcagcc cttggctggt ccctgcgagc 120
ccgtggagac tgccagagat gtcctctttc ggttacagga ccctgactgt ggccctcttc 180
accctgatct gctgtccagg agtcaatgaa ttccaacgtc agcgtgtacc agcctccaag 240
gcaggtcatc ctgacactgc aacccacttt ggtggctgtg ggcaagtcct tcaccattga 300
gtgcagggtg cccaccgtgg agcccctgga cagcctcacc ctcttcctgt tccgtggcaa 360
tgagactctg cactatgaga ccttcgggaa ggcagcccct gctccgcagg aggccacagc 420
cacattcaac agcacggctg acagagagga tggccaccgc aacttctcct gcctggctgt 480
gctggacttg atgtctcgcg gtggcaacat ctttcacaaa cactcagccc cgaagatgtt 540
ggagatctat gagcctgtgt cggacagcca gatggtcatc atagtcacgg tggtgtcggt 600
gttgctgtcc ctgttcgtga catctgtcct gctctgcttc atcttcggcc agcacttgcg 660
ccagcagcgg atgggcacct acggggtgcg agcggcttgg aggaggctgc cccaggcctt 720
ccggccatag caaccatgag tggcatggcc accaccacgg tggtcactgg aactcagtgt 780
gactcctcag ggttgaggtc cagccctggc tgaaggactg tgacaggcag cagagacttg 840
ggacattgcc ttttctagcc cgaatacaaa cacctggact taaaaaaaaa aaaaaaaaa 899

<210> 92
<211> 2529
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510259CB1

<400> 92
gttcatgatg tacgcacctg ttgaattttc agaagctgaa ttctcacgag ctgaatatca 60
aagaaagcag caattttggg actcagtacg gctagctctt ttcacattag caattgtagc 120
aatcatagga attgcaattg gtattgttac tcattttgtt gttgaggatg ataagtcttt 180
ctattacctt gcctctttta aagtcacaaa tatcaaatat aaagaaaatt atggcataag 240
atcttcaaga gagtttatag aaaggagtca tcagattgaa agaatgagat gactaacaag 300
tgagggtgtc agagtcacat ggcagaagtc atttcaagtc atgtcaaatg gtctagtgat 360
gacataatca tttaagaaat caacaaattc tatcaatact ggatgagatg ccatatttga 420
gatgagtgca gcaaacaaga tgtctaggat atttcgacat tcttctgtag gcggtcgatt 480
tatcaaatct catgttatca aattaagtcc agatgaacaa ggtgtggata ttcttatagt 540
gctcatattt cgatacccat ctactgatag tgctgaacaa atcaagaaaa aaattgaaaa 600
ggctttatat caaagtttga agaccaaaca attgtctttg accataaaca aaccatcatt 660
tagactcaca cgctgtggaa taaggatgac atcttcaaac atgccattac cagcatcctc 720
ttctactcaa agaattgtcc aaggaaggga aacagctatg gaaggggaat ggccatggca 780
ggccagcctc cagctcatag ggtcaggcca tcagtgtgga gccagcctca tcagtaacac 840
atggctgctc acagcagctc actgcttttg gaaaaataaa gacccaactc aatggattgc 900
tacttttggt gcaactataa caccacccgc agtgaaacga aatgtgagga aaattattct 960
tcatgagaat taccatagag aaacaaatga aaatgacatt gctttggttc agctctctac 1020
tggagttgag ttttcaaata tagtccagag agtttgcctc ccagactcat ctataaagtt 1080
gccacctaaa acaagtgtgt tcgtcacagg atttggatcc attgtagatg atggacctat 1140
acaaaataca cttcggcaag ccagagtgga aaccataagc actgatgtgt gtaacagaaa 1200
ggatgtgtat gatggcctga taactccagg aatgttatgt gctggattca tggaaggaaa 1260

aatagatgca tgtaagggag attctggtgg acctctggtt tatgataatc atgacatctg 1320
gtacattgta ggtatagtaa gttggggaca atcgtgtgca cttcccaaaa aacctggagt 1380
ctacaccaga gtaactaagt atcgagattg gattgcctca aagactggta tgtagtgtgg 1440
attgtccatg agttatacac atggcacaca gagctggtac tcctgcgtat tttgtattgt 1500
ttaaattcat ttactttgga ttagtgcttt tgctagatgt caagaagccc ttcagaccca 1560
gacaaatcta atatcctgag gtggccttta catacgtagg accaaacccc ctctaccatg 1620
agggaagaag acacagcaaa tgacagacag cacctattcc ttactcacaa gggaaactgc 1680
ttgtgatact tcctaataag ataaataagt ggtttccctc aattgaagac aggaacatca 1740
ttttccacag gatatgaaga gctgccagta atgccaaaat cttacctcat ataatacctg 1800
gagcatgtga gattcttcta gtgaaaaaga acagtcttcc ctgaagactc agggcttcaa 1860
cattctagaa ctgataagtg gaccttcagt gtgcaagaat ggagaagcat gggatttgca 1920
ttatgacttg aactgggctt atatctaata atacagagca ctatcactaa cctcaacagt 1980
tgacatttta aaagttttta aatgtatctg aacttgctgt taacacagtg ttataactca 2040
agcactagct tcaggaagca tgttgtgttg ttaagaagct tttctgattt attctttaac 2100
agcatcttgc catctatatg ttagtagcag ttggcccaga aaggacgaaa aaaagattaa 2160
gactctttgg aacgtttttc catgagcaca ggaggataaa aagaagcaga tgaaggctag 2220
gagaattggt ttcaaataat tagtaacagg acaagcacgc taatttttga tggaatgagt 2280
tatccaatta tttacttaga aatatttata tcagtatatg gcaactggta cttttgtaag 2340
tcttcagctt tctgacaagt cagatgtcca tcagagtatc aggtcaggtg tctatcagaa 2400
tatcagagct gatttgtgta aagctttgtg taaagcacgt aggacagtgc cttgcatata 2460
ctacgaacta aataaatctt gttatatgga aatcaaaaaa aaaaaaaaaa aaaagggggc 2520
cgcccaaga 2529

<210> 93
<211> 2268
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510444CB1

<400> 93
cctacctctt cctaggccca cagccagtgc ctttggagta ctgaggcgcg cacagagtcc 60
ttagcccggc gcagggcgcg cagcccaggc tgagatccgc tgcttctgtg gaagtgagca 120
tggttgggca gcgggtgctg cttctagtgg ccttccttct ttctggggtc ctgctctcag 180
aggctgccaa aatcctgaca atatctacac tgggtgagtg cttggccgga gaattcccag 240
acaggcgcgt cccggatccc cgcactgcca gggctccagc gaacggcgat tgatcagagt 300
tatccaggcg attttccagg ctgggcttgc ggacctggct ggaggaggga gaagcccatc 360
tagccgtggg gcagagaggg gcctctattg ctgaggtgga agccattacc tactgttgga 420
ccgggtgtct cagattcttc aagagcatgg tcataatgtg actatgcttc atcagagtgg 480
aaagtttttg atcccagata ttaaagagga ggaaaaatca taccaagtta tcaggtggtt 540
ttcacctgaa gatcatcaaa aaagaattaa gaagcatttt gatagctaca tagaaacagc 600
attggatggc agaaaagaat ctgaagccct tgtaaagcta atggaaatat ttgggactca 660
atgtagttat ttgctaagca gaaaggatat aatggattcc ttaaagaatg agaactatga 720
tctggtattt gttgaagcat ttgatttctg ttctttcctg attgctgaga agcttgtgaa 780
accatttgtg gccattcttc ccaccacatt cggctctttg gattttgggc taccaagccc 840
cttgtcttat gttccagtat tcccttcctt gctgactgat cacatggact tctggggccg 900
agtgaagaat ttcctgatgt tctttagttt ctccaggagc caatgggaca tgcagtctac 960
atttgacaac accatcaagg agcatttccc agaaggctct aggccagttt tgtctcatct 1020
tctactgaaa gcagagttgt ggtttgttaa ctctgatttt gcctttgatt ttgcccggcc 1080
cctgcttccc aacactgttt atattggagg cttgatggaa aaacctatta aaccagtacc 1140
acaagacttg gacaacttca ttgccaactt tggggatgca gggtttgtcc ttgtggcctt 1200
tggctccatg ttgaacaccc atcagtccca ggaagtcctc aagaagatgc acaatgcctt 1260
tgcccacctc cctcaaggag tgatatggac atgtcagagt tctcattggc ccagagatgt 1320
tcatttggcc acaaatgtga aaattgtgga ctggcttcct cagagtgacc tcctggctca 1380

ccccagcatc cgtctttttg tcactcatgg tgggcagaac agcgtaatgg aggccatccg 1440
tcatggtgtg cccatggtgg gattaccagt caatggagac cagcatggaa acatggtccg 1500
agtagtagcc aaaaattatg gtgtctctat ccggttgaat caggtcacag ccgacacact 1560
gacacttaca atgaaacaag tcatagaaga caagaggtac aagtcggcag tggtggcagc 1620
cagtgtcatc ctgcactctc agcccctgag ccccgcacag cggctggtgg gctggatcga 1680
ccacatcctc cagactgggg gagcgacgca cctcaagccc tatgccttcc agcagccttg 1740
gcatgagcag tacctcattg atgtctttgt gtttctgctg ggctcactc tgggcactat 1800
gtggctttgt gggaagctgc tgggtgtggt ggccaggtgg ctgcgtgggg ccaggaaggt 1860
gaagaagaca tgaggctagg tgtagccttg ggtgagggga gggcatccct ggtcctttga 1920
aggttctccc caccccagca cacgccaccc ctctgttctc tcttcagctc cacctgccac 1980
tgatcctgca acttgcttct ttctattctc tgcctctgtt tagaaatctt cacgcaccac 2040
tgaggcttct tgacttgccc cttgtgactt gaaaccccag ctcagataca aattttcacc 2100
tgccagccct gcctcctcct ttctcccttt tcctagacac aggactctga caacttcatc 2160
ctccttgttt agatgacttc ccagtttcca gtccccattt ctccttctat cacttttcat 2220
aaaaaaactc aggaaatatt tgacatatct tccatttcaa attcttcc 2268

<210> 94
<211> 1819
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510494CB1

<400> 94
ctcttctttg cctctgctgg agtccgggga gtggcgttgg ctgctagagc gatgccgggc 60
cggagttgcg tcgccttagt cctcctggct gccgccgtca gctgtgccgt cgcgcagcac 120
gcgccgccgg tgagtgagct tgagccgagg cgcagagagg ggcgtgcagt ggacagagga 180
ctgcagaaaa tcaacctatc ctccttcagg accaacgtac agaggtgcag ttccatggta 240
caccataaat cttgacttac caccctacaa aagatggcat gaattgatgc ttgacaaggc 300
accaatgcta aaggttatag tgaattctct gaagaatatg ataaatacat tcgtgccaag 360
tggaaaagtt atgcaggtgg tggatgaaaa attgcctggc ctacttggca actttcctgg 420
ccctttgaa gaggaaatga agggtattgc cgctgttact gatatacctt taggagagat 480
tatttcattc aatattttt atgaattatt taccatttgt acttcaatag tagcagaaga 540
caaaaaaggt catctaatac atgggagaaa catggatttt ggagtatttc ttgggtggaa 600
cataaataat gatacctggg tcataactga gcaactaaaa cctttaacag tgaatttgga 660
tttccaaaga aacaacaaaa ctgtcttcaa ggcttcaagc tttgctggct atgtgggcat 720
gttaacagga ttcaaaccag gactgttcag tcttacactg aatgaacgtt tcagtataaa 780
tggtggttat ctgggtattc tagaatggat tctgggaaag aaagatgcca tgtggatagg 840
gttcctcact agaacagttc tggaaaatag cacaagttat gaagaagcca agaatttatt 900
gaccaagacc aagatattgg ccccagccta ctttatcctg ggaggcaacc agtctgggga 960
aggttgtgtg attacacgag acagaaagga atcattggat gtatatgaac tcgatgctaa 1020
gcagggtaga tggtatgtgg tacaaacaaa ttatgaccgt tggaaacatc ccttcttcct 1080
tgatgatcgc agaacgcctg caaagatgtg tctgaaccgc accagccaag agaatatctc 1140
atttgaaacc atgtatgatg tcctgtcaac aaaacctgtc ctcaacaagc tgaccgtata 1200
cacaaccttg atagatgtta ccaaaggtca attcgaaact tacctgcggg actgccctga 1260
cccttgtata ggttggtgag cacacgtctg gcctacagaa tgcggcctct gagacatgaa 1320
gacaccatct ccatgtgacc gaacactgca gctgtctgac cttccaaaga ctaagactcg 1380
cggcaggttc tctttgagtc aaaagcttgt cttcgtccat ctgttgacaa atgacagatc 1440
ttttttttt ccccctatca gttgattttt cttatttaca gataacttct ttaggggaag 1500
taaaacagtc atctagaatt cactgagttt ggtttcactt tgacatttgg ggatctggtg 1560
ggcagtcgaa ccatggtgaa ctccacctcc gtggaataaa tggagattca gcgtgggtgt 1620
tgaatccagc acgtctgtgt gagtaacggg acagtaaaca ctccacattc ttcagttttt 1680
cacttctacc tacatatttg tatgtttttc tgtataacag ccttttcctt ctggttctaa 1740
ctgctgttaa aattaatata tcattatctt tgctgttatt gacagcgata taattttatt 1800

acatatgatt agagggatg 1819

<210> 95
<211> 1711
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6486485CB1

<400> 95
aaactatttg agggtagggg ctgtgattat ttactctcat atcctcagag cctggtgttg 60
aggttggtgc tttgtaggca cccagggact ttcaaatgaa tgaagggagg gagggaggaa 120
agaaggatgg gtccatagta ggacctggtg atgggctggg agctccaggc aaatgtcaac 180
caatccctct cctgggtcag ctcccagggg ctcacccttc tttgcatttc cagctctcat 240
gaggtcattg tgcacaggaa agctctctcc tctaatctcc tctgatccta ctgcaccaga 300
gaaatcaagc cagaattcaa caaagtctca gtccagataa acaagacaaa agaaataaga 360
ttcgagtaga agatctcctt caagggaaag ttgctgtgtt tgtccaagac ctttgtccca 420
tccatgtatc atcccccaag taaacacttc ttgttcacct gttcattaga tttcaagtgc 480
agtccctggc ctgtaagtcc ctacaatgat aagtttctct tatcattgca cattcttcat 540
caggaggatg ccagaggagc tcagccaaca gttcctcatc agtagcagat tcttcagaat 600
cttgggcact acacagatgc ccttgagctc tttgaataaa ggctgatttt tagaaaaaac 660
attaagacag aacttaaaaa caatagattg actataatcc aaagacgagt gtacctctaa 720
ccacaatttt catttatttt taaatgtttc cttcatggcc tttcttgtgg ctcaccctat 780
gcagtttgtg tatttgttga caactttatg tgtttttaat atggtttttg ccaaacttgg 840
tttttccgag accgtctttt ctcagaggct cagttttacc gtcctatctg cagtcggcta 900
ctttcagtgg cagaagaggc cacatctgct tcctgtaggc cctctgggca gaagcatgcg 960
ctggtgtctc ctcctgatct gggcccaggg gctgaggcag gctcccctcg cctcaggaat 1020
gatgacaggc acaatagaaa caacggggaa catttctgca gagaaaggtg gctctatcat 1080
cttacaatgt cacctctcct ccaccacggc acaagtgacc caggtcaact gggagcagca 1140
ggaccagctt ctggccattt gtaatgctga cttggggtgg cacatctccc catccttcaa 1200
ggatcgagtg gccccaggtc ccggcctggg cctcaccctc cagtcgctga ccgtgaacga 1260
tacaggggag tacttctgca tctatcacac ctaccctgat gggacgtaca ctgggagaat 1320
cttcctggag gtcctagaaa gctcagtggc tgagcacggt gccaggttcc agattccatt 1380
gcttggagcc atggccgcga cgctggtggt catctgcaca gcagtcatcg tggtggtcgc 1440
gttgactaga aagaagaaag ccctcagaat ccattctgtg gaaggtgacc tcaggagaaa 1500
atcagctgga caggaggaat ggagccccag tgctccctca cccccaggaa gctgtgtcca 1560
ggcagaagct gcacctgctg ggctctgtgg agagcagcgg ggagaggact gtgccgagct 1620
gcatgactac ttcaatgtcc tgagttacag aagcctgggt aactgcagct tcttcacaga 1680
gactggttag caaccagagg catcttctgg a 1711

<210> 96
<211> 5096
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503772CB1

<400> 96
gggaacccag aaccagcccg agccgcctgc cccgtcgccc ggccgccggc ttagggcgca 60
gcgcggttgg tcctcgcccc ctcccgcccg ccggcctacc aggccatggg ggcgtcccgg 120
gaccgcgggc tggccgcgct ctggtgcctt gggctcctgg ggggcctggc gcgcgtcgcg 180
ggcacgcact accgctacct ctggaggggc tgctacccat gtcacctggg ccaggccggc 240

taccccgtga gcgccggtga ccagaggcca gatgtggacg aatgccgaac ccacaacggt 300
ggctgccagc accggtgcgt gaacacccca ggctcctacc tctgtgagtg caagcccggc 360
ttccggctcc acactgacag caggacctgc ctggccatta actcctgcgc cctgggcaat 420
ggcggctgcc agcaccactg tgtccagctc acaatcactc ggcatcgctg ccagtgccgg 480
cccgggttcc agctccagga ggacggcagg cattgtgtcc gtagaagccc gtgtgccaac 540
aggaacggca gctgcatgca caggtgccag gtggtccggg gcctcgcccg ctgtgagtgc 600
cacgtgggct atcagctagc agcggacggc aaggcctgtg aagatgtgga cgaatgtgcc 660
gcagggctgg cccagtgtgc ccatggctgc ctcaacaccc aggggtcctt caagtgcgtg 720
tgtcacgcgg gctatgagct gggagccgat ggccggcagt gctaccggat tgagatggaa 780
atcgtgaaca gctgtgaggc caacaacggc ggctgctccc atggctgcag ccacaccagt 840
gctgggcccc tgtgcacatg tccccgcggc tacgagctgg acacagatca gaggacctgc 900
atcgatgtcg acgactgtgc agacagcccg tgctgccagc aggtgtgcac caacaacccct 960
ggcgggtacg agtgcggctg ctacgccggc taccggctca gtgccgatgg ctgcggctgt 1020
gaggatgtgg atgagtgcgc ctccagccgt ggcggctgcg agcaccactg caccaacctg 1080
gccggctcct tccagtgctc ctgcgaggcc ggctaccggc tgcacgagga ccgtaggggc 1140
tgcagccccc tggaggagcc gatggtggac ctggacggcg agctgccttt cgtgcggccc 1200
ctgccccaca ttgccgtgct ccaggacgag ctgccgcaac tcttccagga tgacgacgtc 1260
ggggccgatg aggaagaggc agagttgcgg ggcgaacaca cgctcacaga gaagtttgtc 1320
tgcctggatg actcctttgg ccatgactgc agcttgacct gtgatgactg caggaacgga 1380
gggacctgcc tcctgggcct ggatggctgt gattgccccg agggctggac tgggctcatc 1440
tgcaatgaga cttgtcctcc ggacaccttt gggaagaact gcagcttctc ctgcagctgt 1500
cagaatggtg ggacctgcga ctctgtcacg gggggcctgcc gctgccccc gggtgtcagt 1560
ggaactaact gtgaggatgg ctgccccaag ggctactatg gcaagcactg tcgcaagaaa 1620
tgcaactgtg ccaaccgggg ccggtgccac cgcctctacg ggcctgcct ctgcgaccca 1680
gggctctacg gccgcttctg ccacctcacc tgcccgccgt gggcctttgg gccgggctgc 1740
tcggaggagt gccagtgtgt gcagccccac acgcagtcct gtgacaagag ggatggcagc 1800
tgctcctgca aggctggctt ccggggcgag cgctgtcagg cagagtgtga gctgggctac 1860
tttgggccgg ggtgctggca ggcatgcacc tgcccagtgg gcgtggcctg tgactccgtg 1920
agcggcgagt gtgggaagcg gtgtcctgct ggcttccagg gagaggactg tggccaagag 1980
tgcccggtgg ggacgtttgg cgtgaactgc tcgagctcct gctcctgtgg gggggccccc 2040
tgccacgggg tcacggggca gtgccggtgt ccgccgggga ggactgggga agactgtgag 2100
gcagattgtc ccgagggccg ctgggggctg ggctgccagg agatctgccc agcatgccag 2160
cacgctgccc gctgcgaccc tgagaccgga gcctgcctgt gcctccctgg cttcgtcggc 2220
agccgctgcc aggacgtgtg cccagcaggc tggtatggtc ccagctgcca gacaaggtgc 2280
tcttgtgcca atgatgggca ctgccaccca gccaccggac actgcagctg tgccccccggg 2340
tggaccggct ttagctgcca gagagcctgt gatactgggc actggggacc tgactgcagc 2400
cacccctgca actgcagcgc tggccacggg agctgtgatg ccatcagcgg cctgtgtctg 2460
tgtgaggctg gctacgtggg cccgcggtgc gagcagcagt gtccccaggg ccactttggg 2520
cccggctgtg agcagctgtg ccagtgtcag catggagcag cctgtgacca cgtcagcggg 2580
gcctgcacct gcccggccgg ctggaggggc accttctgcg agcatgcctg cccggccggc 2640
ttctttggat tggactgtcg cagtgcctgc aactgcaccg ccggagctgc ctgtgatgcc 2700
gtgaatggct cctgcctctg ccccgctggc cgccggggcc cccgctgtgc cgagacctgc 2760
ccagcccaca cctacgggca caattgcagc caggcctgtg cctgctttaa cggggcctcc 2820
tgtgaccctg tccacgggca gtgccactgt gcccctggct ggatggggcc ctcctgcctg 2880
caggagtgcc tcccccggga cgtcagagct ggctgccggc acagcggcgg ttgcctcaac 2940
gggggcctgt gtgacccgca cacgggccgc tgcctctgcc cagccggctg gactggggac 3000
aagtgtcaga gcccctgcct gcggggctgg tttggagagg cctgtgccca gcgctgcagc 3060
tgcccgcctg gcgctgcctg ccaccacgtc actggggcct gccgctgtcc ccctggcttc 3120
actggctccg gctgcgagca gggatgtccg cccgggcggt atgggccagg ctgtgaacag 3180
ctgtgtgggt gtctcaacgg gggctcctgt gatgcggcca cgggggcctg ccgctgcccc 3240
actgggttcc tcgggacgga ctgcaacctc agtgagtggc tggtggcagc agtgctgagt 3300
ggcgcctcct gtgtctccag gcctgtccgc agggccgctt cggccccaac tgcacccacg 3360
tgtgtgggtg tgggcagggg gcggcctgcg accctgtgac cggcacctgc ctctgccccc 3420
cggggagagc cggcgtccgc tgtgagcgag gctgccccca gaaccggttt ggcgtgggct 3480
gcgagcacac ctgctcctgc agaaatgggg gcctgtgcca cgccagcaac ggcagctgct 3540
cctgtggcct gggctggacg ggcggcact gcgagctggc ctgtcccccct gggcgctacg 3600

gagccgcctg ccatctggag tgctcctgcc acaacaacag cacgtgtgag cctgccacgg 3660
gcacctgccg ctgcggcccc ggcttctatg gccaggcctg cgagcacccc tgtccccctg 3720
gcttccacgg ggctggctgc caggggttgt gctggtgtca acatggagcc ccctgcgacc 3780
ccatcagtgg ccgatgcctc tgccctgccg gcttccacgg ccacttctgt gagagggggt 3840
gtgagccagg ttcatttgga gagggctgcc accagcgctg tgactgtgac ggggggcac 3900
cctgtgaccc tgtcaccggt ctctgccttt gcccaccagg gcgctcagga gccacctgta 3960
acctggattg cagaagggc cagtttgggc ccagctgcac cctgcactgt gactgcgggg 4020
gtggggctga ctgcgaccct gtcagtgggc agtgtcactg tgtggatggc tacatggggc 4080
ccacgtgccg ggaaggtggg cccctccggc tccccgagaa cccgtcctta gcccagggct 4140
cagcgggcac actgcccgcc tccagcagac ccacatcccg gagcggtgga ccagcgaggc 4200
actagtagag gcagtcccgt ggagcccgcc tctccagtcc cagccagagg ggaccctggc 4260
ctttggtgac cactgagaag gacacttcac gggcccagag ctcctggtac tgcccttcct 4320
ttgagggccg tggagggctg tggacagccc agcaacctgt cgctcttgga ggctggtgtg 4380
gccttgagga gggaagcctc gcatggccgc tggaagagag gcgcctcctg gcctggctct 4440
gcagaaccca ggggcacgct ctgggcctgg gctgaggaag tcccgctctc cccgcggctc 4500
tgagtcggac tgaggacagg tgtgggcgcc agtgtgggtg caggcgcagg tgcaggcaca 4560
gggccactgt cctccaggca ggctttttgg tgctaggccc tgggactgga agtcgcccag 4620
cccgtattta tgtaaaggta tttatgggcc actgcacatg cccgctgcag ccctgggatc 4680
agctggaagc tgcctgtcat ctcctgccca atccccagaa accctgattc aggtctgcag 4740
gctcctgcgg gctcaccagg ctgctggctc cggtaccatg taaacctagg aaggtaaagg 4800
agcaggcaac ctcctcgtgg cctgtgtgtt tgctgtgtta cgtggactct gtgtgggctc 4860
ctccctgggg cccggccagc ataacggtgc acccagggac ctcccagtgc acccggggcc 4920
ctttgcaggg gtgggggtgc cacacaagtg aagaagttgg gactcatctc agttcccagt 4980
gctattgagg agaacgctgg ggctgcattc attaccgctg agacccagag actggctgtt 5040
cccagagaat ggcccagggg gaggagggct ggtgtggaag ggcaacttgg actgag 5096

<210> 97
<211> 5049
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503773CB1

<400> 97
gggaacccag aaccagcccg agccgcctgc cccgtcgccc ggccgccggc ttagggcgca 60
gcgcggttgg tcctcgcccc ctcccgcccg ccggcctacc aggccatggg ggcgtcccgg 120
gaccgcgggc tggccgcgct ctggtgcctt gggctcctgg ggggcctggc gcgcgtcgcg 180
ggcacgcact accgctacct ctggaggggc tgctacccat gtcacctggg ccaggccggc 240
taccccgtga gcgccggtga ccagaggcca gatgtggacg aatgccgaac ccacaacggt 300
ggctgccagc accggtgcgt gaacacccca ggctcctacc tctgtgagtg caagcccggc 360
ttccggctcc acactgacag caggacctgc ctggccatta actcctgcgc cctgggcaat 420
ggcggctgcc agcaccactg tgtccagctc acaatcactc ggcatcgctg ccagtgccgg 480
cccgggttcc agctccagga ggacggcagg cattgtgtcc gtagaagccc gtgtgccaac 540
aggaacggca gctgcatgca caggtgccag gtggtccggg gcctcgcccg ctgtgagtgc 600
cacgtgggct atcagctagc agcggacggc aaggcctgtg aagatgtgga cgaatgtgcc 660
gcagggctgg cccagtgtgc ccatggctgc ctcaacaccc aggggtcctt caagtgcgtg 720
tgtcacgcgg gctatgagct gggagccgat ggccggcagt gctaccggat tgagatggaa 780
atcgtgaaca gctgtgaggc caacaacggc ggctgctccc atggctgcag ccacaccagt 840
gctgggcccc tgtgcacatg tccccgcggc tacgagctgg acacagatca gaggacctgc 900
atcgatgtcg acgactgtgc agacagcccg tgctgccagc aggtgtgcac caacaaccct 960
ggcgggtacg agtgcggctg ctacgccggc taccggctca gtgccgatgg ctgcggctgt 1020
gaggatgtgg atgagtgcgc ctccagccgt ggcggctgcg agcaccactg caccaacctg 1080
gccggctcct tccagtgctc ctgcgaggcc ggctaccggc tgcacgagga ccgtaggggc 1140
tgcagccccc tggaggagcc gatggtggac ctggacggcg agctgccttt cgtgcggccc 1200

ctgccccaca ttgccgtgct ccaggacgag ctgccgcaac tcttccagga tgacgacgtc 1260
ggggccgatg aggaagaggc agagttgcgg ggcgaacaca cgctcacaga gaagtttgtc 1320
tgcctggatg actcctttgg ccatgactgc agcttgacct gtgatgactg caggaacgga 1380
gggacctgcc tcctgggcct ggatggctgt gattgccccg agggctggac tgggctcatc 1440
tgcaatgaga cttgtcctcc ggacaccttt gggaagaact gcagcttctc ctgcagctgt 1500
cagaatggtg ggacctgcga ctctgtcacg gggggcctgcc gctgccccc gggtgtcagt 1560
ggaactaact gtgaggatgg ctgccccaag ggctactatg gcaagcactg tcgcaagaaa 1620
tgcaactgtg ccaaccgggg ccggtgccac cgcctctacg ggcctgcct ctgcgaccca 1680
gggctctacg gccgcttctg ccacctcacc tgcccgccgt gggcctttgg gccgggctgc 1740
tcggaggagt gccagtgtgt gcagccccac acgcagtcct gtgacaagag ggatggcagc 1800
tgctcctgca aggctggctt ccggggcgag cgctgtcagg cagagtgtga gctgggctac 1860
tttgggccgg ggtgctggca ggcatgcacc tgcccagtgg gcgtggcctg tgactccgtg 1920
agcggcgagt gtgggaagcg gtgtcctgct ggcttccagg gagaggactg tggccaagag 1980
tgcccggtgg ggacgtttgg cgtgaactgc tcgagctcct gctcctgtgg gggggccccc 2040
tgccacgggg tcacggggca gtgccggtgt ccgccgggga ggactgggga agactgtgag 2100
gcagattgtc ccgagggccg ctgggggctg ggctgccagg agatctgccc agcatgccag 2160
cacgctgccc gctgcgaccc tgagaccgga gcctgcctgt gcctccctgg cttcgtcggc 2220
agccgctgcc aggacgtgtg cccagcaggc tggtatggtc ccagctgcca gacaaggtgc 2280
tcttgtgcca atgatgggca ctgccaccca gccaccggac actgcagctg tgcccccggg 2340
tggaccggct ttagctgcca gagagcctgt gatactgggc actggggacc tgactgcagc 2400
cacccctgca actgcagcgc tggccacggg agctgtgatg ccatcagcgg cctgtgtctg 2460
tgtgaggctg gctacgtggg cccgcggtgc gagcagcagt gtccccaggg ccactttggg 2520
cccggctgtg agcagctgtg ccagtgtcag catggagcag cctgtgacca cgtcagcggg 2580
gcctgcacct gcccggccgg ctggaggggc accttctgcg agcatgcctg cccggccggc 2640
ttctttggat tggactgtcg cagtgcctgc aactgcaccg ccggagctgc ctgtgatgcc 2700
gtgaatggct cctgcctctg ccccgctggc cgccggggcc cccgctgtgc cgagacctgc 2760
ccagcccaca cctacgggca caattgcagc caggcctgtg cctgctttaa cggggcctcc 2820
tgtgaccctg tccacgggca gtgccactgt gcccctggct ggatggggcc ctcctgcctg 2880
caggagtgcc tcccccggga cgtcagagct ggctgccggc acagcggcgg ttgcctcaac 2940
gggggcctgt gtgacccgca cacgggccgc tgcctctgcc cagccggctg gactggggac 3000
aagtgtcaga gcccctgcct gcggggctgg tttggagagg cctgtgccca gcgctgcagc 3060
tgcccgcctg gcgctgcctg ccaccacgtc actggggcct gccgctgtcc ccctggcttc 3120
actggctccg gctgcgagca ggcctgccca cccggcagct tggggagga ctgtgcgcag 3180
atgtgccagt gtcccggtga gaacccggcc tgccaccctg ccaccgggac ctgctcatgt 3240
gctgctggct accacggccc cagctgccag caacgatgtc cgcccgggcg gtatgggcca 3300
ggctgtgaac agctgtgtgg gtgtctcaac gggggctcct gtgatgcggc cacgggggcc 3360
tgccgctgcc ccactgggtt cctcgggacg gactgcaacc tcacctgtcc gcagggccgc 3420
ttcggcccca actgcaccca cgtgtgtggg tgtgggcagg gggcggcctg cgaccctgtg 3480
accggcacct gcctctgccc cccggggaga gccggcgtcc gctgtgagcg aggctgcccc 3540
cagaaccggt ttggcgtggg ctgcgagcac acctgctcct gcagaaatgg gggcctgtgc 3600
cacgccagca acggcagctg ctcctgtggc ctgggctgga cggggcggca ctgcgagctg 3660
gcctgtcccc ctgggcgcta cggagccgcc tgccatctgg agtgctcctg ccacaacaac 3720
agcacgtgtg agcctgccac gggcacctgc cgctgcggcc ccggcttcta tggccaggcc 3780
tgcgagcacc cctgtccccc tggcttccac ggggctggct gccaggggtt gtgctggtgt 3840
caacatggag cccccctgcga ccccatcagt ggccgatgcc tctgccctgc cggcttccac 3900
ggccacttct gtgagagggg gtgtgagcca ggttcatttg gagagggctg ccaccagcgc 3960
tgtgactgtg acgggggggc accctgtgac cctgtcaccg gtctctgcct ttgcccacca 4020
gggcgctcag gagccacctg taacctgggt gggcccctcc ggctccccga gaacccgtcc 4080
ttagcccagg gctcagcggg cacactgccc gcctccagca gacccacatc ccggagcggt 4140
ggaccagcga ggcactagta gaggcagtcc cgtggagccc gcctctccag tcccagccag 4200
aggggaccct ggcctttggt gaccactgag aaggacactt cacgggccca gagctcctgg 4260
tactgccctt cctttgaggg ccgtggaggg ctgtggacag cccagcaacc tgtcgctctt 4320
ggaggctggt gtggccttga ggagggaagc ctcgcatggc cgctggaaga gaggcgcctc 4380
ctggcctggc tctgcagaac ccaggggcac gctctgggcc tgggctgagg aagtcccgct 4440
ctccccgcgg ctctgagttg gactgaggac aggtgtgggc gccagtgtgg gtgcaggcgc 4500
aggtgcaggc acagggccac tgtcctccag gcaggctttt tggtgctagg ccctgggact 4560

ggaagtcgcc cagcccgtat ttatgtaaag gtatttatgg gccactgcac atgcccgctg 4620
cagccctggg atcagctgga agctgcctgt catctcctgc ccaatcccca gaaaccctga 4680
ttcaggtctg caggctcctg cgggctcacc aggctgctgg ctccggtacc atgtaaacct 4740
aggaaggtaa aggagcaggc aacctcctcg tggcctgtgt gtttgctgtg ttacgtggac 4800
tctgtgtggg ctcctccctg gggcccggcc agcataacgg tgcacccagg gacctcccag 4860
tgcacccggg gccctttgca ggggtggggg tgccacacaa gtgaagaagt tgggactcat 4920
ctcagttccc agtgctattg aggagaacgc tggggctgca ttcattaccg ctgagaccca 4980
gagactggct gttcccagag aatggcccag gggaggaggg gctggtgtgg aagggcaact 5040
tggactgag 5049

<210> 98
<211> 1086
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504698CB1

<400> 98
gcgggacttc ctgtgtcgta tttccaagga ctccaaagcg aggccgggga ctgaaggtgt 60
gggtgtcgag ccctctggca gagggttaac ctgggtcaaa tgcacggatt ctcacctcgt 120
acagttacgc tctcccgcgg cacgtccgcg aggacttgaa gtcctgagcg ctcaagtttg 180
tccgtaggtc gagagaaggc catggaggtg ccgccaccgg caccgcggag ctttctctgt 240
agagcattgt gcctatttcc ccgagtcttt gctgccgaag ctgtgactgc cgattcggaa 300
gtccttgagg agcgtcagaa gcggcttccc tacgtcccag agccctatta cccggaatct 360
ggatgggacc gcctccggga gctgtttggc aaagactgtc acgggaagtc tttttaggat 420
aaacgtaggc ctgcgtggcc tggtggctgg tggcataatt ggagccttgc tgggcactcc 480
tgtaggaggc ctgctgatgg catttcagaa gtactctggt gagactgttc aggaaagaaa 540
acagaaggat cgaaaggcac tccatgagct aaaactggaa gagtggaaag gcagactaca 600
agttactgag cacctccctg agaaaattga aagtagttta caggaagatg aacctgagaa 660
tgatgctaag aaaattgaag cactgctaaa ccttcctaga aacccttcag taatagataa 720
acaagacaag gactgaaagt gctctgaact tgaaactcac tggagagctg aagggagctg 780
ccatgtccga tgaatgccaa cagacaggcc actctttggt cagcctgctg acaaatttaa 840
gtgctggtac ctgtggtggc agtggcttgc tcttgtcttt ttcttttctt tttaactaag 900
aatggggctg ttgtactctc actttactta tccttcaatt taaatacata cttatgtttg 960
tattaatcta tcaatatatg catacatgaa tatatccacc cacctagatt ttaagcagta 1020
aataaaacat ttcgcaaaag attaaagttg aattttacag ttaaaaaaaa aaaaaaaaaa 1080
aaaact 1086

<210> 99
<211> 1099
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510361CB1

<400> 99
caaggatttc atgagcatcc tcctctaaac gcgtgtcaag acaaaagatg cttcagcttt 60
ggaaacttgt tctcctgtgc ggcgtgctca ctgggacctc agagtctctt cttgacaatc 120
ttggcaatga cctaagcaat gtcgtggata agctggaacc tgttcttcac gagggacttg 180
agacagttga caatactctt aaaggcatcc ttgagaaact gaaggtcgac ctaggagtgc 240
ttcagaaatc cagtgcttgg caactggcca agcagaaggc ccaggaagct gagaaattgc 300
tgaacaatgt catttctaag ctgcttccaa ctaacacgga catttttggg ttgaaaatca 360

gcaactccct catcctggat gtcaaagctg aaccgatcga tgatggcaaa ggccttaacc 420
tgagcttccc tgtcaccgcg aatgtcactg tggccgggcc catcattggc cagattatca 480
acctgaaagc ctccttggac ctcctgaccg cagtcacaat tgaaactgat ccccagacac 540
accagcctgt tgccgtcctg ggagaatgcg ccagtgaccc aaccagcatc tcactttcct 600
tgctggacaa ccaaaaatgt attgaggctg ggcacgatgg ctccacaccg gtaatcccag 660
cactttggga ggctgagaca cagccaaatc atcaacaagt tcgtgaatag cgtgatcaac 720
acgctgaaaa gcactgtatc ctccctgctg cagaaggaga tatgtccact gatccgcatc 780
ttcatccact ccctggatgt gaatgtcatt cagcaggtcg tcgataatcc tcagcacaaa 840
acccagctgc aaaccctcat ctgaagagga cgaatgagga ggaccactgt ggtgcatgct 900
ggtgaggagc cagtctctgt gccccaatgc acaggggcct atggtgaagt aaaagtcaag 960
cgtggcttcc cttattttcg tgttagaaga ctgtgccttc atctcagtca tagattgagc 1020
cctagccccc atcccaggct aggccctgat tctggtcaca ctctgaacac tgagccttaa 1080
tcactgaact ctgattctg 1099

<210> 100
<211> 1110
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507013CB1

<400> 100
gcctggttac cgggagtggg gcgcccctcc tccttatccc ctcccctctt ccctgtcccc 60
tttcacagct ggctgtagct ggccaaggag ttctcgatta aagaggaagg ggcagtgctc 120
acatttctgg gcaggtgtct ggaaaatatg aggggccacc cctctctgct gctgctatat 180
atggcattaa ccacctgcct ggatacttca cccagtgagg agacagacca agaggtgttc 240
ctgggaagag gccagggagt attttgagga caacactctc acggagcgct tttgggagag 300
ctacatctac aatggcaaag gagggcgtgg acgagtggat gtggccagcc tggctgtggg 360
gctgacaggt ggcatcctgc tcattgtcct ggccggcctg ggagcctttt ggtatctgcg 420
ctggcgacag caccgaggcc agcagccctg tccccaagag gccgggctca ttagccctct 480
gaggcctcac tgaagagctg ctttcgagac ccggctctcc gaaccgtgcc cctgattcat 540
accggattcc ggaagccgct aggcctcata gacgccgaag ctggacttgg agtggggaat 600
ggtgggagta ggggtcatcc ggcccgaggc ctgccctggc acacgcgttt ccgccgcgta 660
tggatataca catgttttcg gcaacgtgtt cccgtgtcct ggcccctcac gggcccccac 720
actctcctga ccgtgagggc actggtcagt tccgcccccg tggtaggcag acgcgcgggg 780
aaattcggac ccaggagccc agccccggct gtgccatctt gtgtatgggc agatatgacc 840
tgacagcccc ctccagtgcc acagggtacg cacacgcaga gccccgcctg tgcacacgcg 900
tgtcttcgtg cactccccgt gcggtacagg ggcacttcgt aacccaggga aagggcgggg 960
ggcatatttg caagcgcgct cggtgcgggc aggctcgcat tgcacccagg gagctggagt 1020
tgagctgttc ccctaaataa aaaccttcgg aaaggagacc aaaaaaagca gaaataatgc 1080
aaaaaataat aaatgaaatg aactgcgaaa 1110

<210> 101
<211> 3082
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510507CB1

<400> 101
gggagaacgc gagtgttagt taacatggct tggtgtgtgg ccacagctga cccggcgcat 60
acctcccgtc ctctctttac agggctggcg gtgtctcggg gctcggctgg ccatgcctgg 120

agcgcggggt ttgactgggc cgccgttgtt gtagtgaccg ggaggcgctg ccgctctggg 180
cagacggttc cgggagccgc acggtcccct ctccttcccc atcctctccc ctcccctctc 240
cgggttcccc cacccacagg agccttgggc cgaccactcc cccgatggcc tcagccacgg 300
aggaccccgt tctggagcgt tatttcaaag gccacaaagc tgcgatcacc tccttggacc 360
tcagccccaa cggcaagcaa cttgctactg cttcttggga taccttttctc atgctatgga 420
atttcaagcc acatgctaga gcttacagat atgtgggtca caaggatgtt gtaaccagcg 480
tgcagttttc tccacatgga aacttattgg cgtctgcctc acgagacaga accgtgagac 540
tctggattcc tgataagaga ggaaaattct cagaatttaa agctcataca gctccagttc 600
gaagtgtaga cttttcagct gatggccagt ttctagctac agcttctgaa gacaaatcca 660
taaaagtatg gagcatgtat cgccagcgct tcctgtattc cttgtatcga catacacact 720
gggtacgctg tgccaaattt tcacccgatg gaagactaat tgtgtcatgt agtgaggata 780
aaactattaa aatttgggat accacaaata agcaatgtgt taataacttc tcagatttcg 840
ttggatttgc aaattttgtg gactttaacc ctagtggtac atgcatagct tcagcaggtt 900
ctgatcaaac tgtgaaagtc tgggatgtaa gagtgaacaa attactacag cattaccaag 960
ttcacagcgg tggagttaat tgcatatcat tccatccttc gggtaactat ctcatcacag 1020
cttcttcaga tggtaccctt aagattctgg acctcttaga aggaaggctc atctatacac 1080
ttcaaggaca tacgggacct gtctttactg tttcattttc aaaaggtgga gagctatttg 1140
catcaggagg tgcagacaca caggtcttat tatggaggac taactttgat gaattgcatt 1200
gtaaaggtct taccaaaaga aatctcaaaa gattacattt tgattcacca ccacatcttc 1260
ttgatatcta cccaagaaca ccacatcccc atgaggaaaa agttgagact gtagaaacaa 1320
cagaaaccag tggtaggact ctgccagaca agggtgaaga ggcctgtgga tatttcttga 1380
acccttcctt aatgtcacca gaatgtttgc caacaaccac gaaaaagaaa acagaagaca 1440
tgagtgacct cccctgtgaa agtcaaagga gcatacctct cgctgtgact gatgctttag 1500
agcatattat ggaacaactc aatgttttga cacagactgt ttcaatcttg gagcagcgac 1560
tgactttgac agaggataag ctgaaagact gccttgaaaa tcagcaaaag cttttcagtg 1620
ctgtccaaca gaaaagctga ataaaaaatt catttcatt tgttgggcag aggcccaata 1680
aatgaacaaa tgtacataca ctcaggaagg tagtacaaga tactccatac aacacaacca 1740
tgtgctattt atcatggcat ttcttaaaag ggtgagcaac agaacaaaag gcagaaaagg 1800
catacctaag gactaattta aacacatatc aatgtgaagg actaatttaa attactatca 1860
tttatgattg cagtaataaa gtgataagca ttcaagcaac tctgtatttt ccccatatta 1920
ttttaaatgt ccattttcat ttataggcca aatcctgcca ggaaagtaac cagatctctg 1980
gatttcactg ttaagtcatt tcagattgac catattcaga cagtcatggg gtgaaataat 2040
tcacttacct ccaaaatagc atcctatatg ccaataatga gttattgatc tgactagttg 2100
tatgtctttc tgttcaaaat agaaattatc ctttcttact aatgccttga aagaatgaac 2160
aaataaaaat tcccagacca cagaatttcc acagcaagaa tacacttatt ttaattaaca 2220
atagcacaga tatagcatag ggcagtgggt tttttagtta atttatggcg tactttgttt 2280
atccattggc caacctgaag gaaatgaaac tcacctatct ttctatcaca gatgaatgtg 2340
ctagatgaat gatttgggtt gtatctgatc atggttcaca aaaattatgt tagtgtgttt 2400
tcagtatgct aaaaagtcag agtgatacaa aagtgatatt taaaaatata cacacatgta 2460
tgtacacatg cttcagaaat attgttgctt agggtgattt gggcagctaa atagtaagta 2520
ctttttaaaa tttttgcatc atcatcttcc tatttaatga attgtgattt aaaacaaaat 2580
gaaaataagc caggtattct aaaagatcct ggatacaaat taagaatttt gctttatttt 2640
aaaccaaatt gagattaaat tgaagaaaag caagcaaatt aatttcagct tgattatcaa 2700
cctgtatcaa gaacaaaaat gggaggaggt gtccacattt atggtgtgta taggtaacat 2760
ggggaaaatg ctattctgtg ttttggaaaa gaagaaatag tgccgtccta tttatttcta 2820
tatttagaaa tttttctcaa agaaatttca attgtatcta tgagatgggt ttctaagtat 2880
cttattgtgt gttataagtg cctttttaata tcatactaag tgtgagcttc tggacatttt 2940
caagagctta caaaaactaa gtggcattgt atttttataa ccccattgag aagactaagt 3000
aagaaatgaa atgtcctatc aattttattt tgtcatgctt caaacaataa agacatttct 3060
gctttaaaaa aaaaaaaaaa gg 3082

<210> 102
<211> 1548
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 90106370CB1

<400> 102
cctctgcctc atttacttcc tcttctcacc cctgtgtctc tcagatgccg ttgactccag 60
agccgccctc tgggcgcgtg gaggggcccc ccgcatggga agcagcccca tggccctcac 120
tgccctgtgg gccctgcatc cccatcatgc tggtcctggc caccctggct gcgctcttca 180
tcctcaccac cgctgtgttg gctgaacgcc tgttccgccg tgctctccgc ccagacccca 240
gccaccgtgc acccaccctg gtgtggcgcc caggaggaga gctgtggatt gagcccatgg 300
gcaccgcccg agagcgctct gaggactggt atggctctgc ggtcccccctg ctgacagatc 360
gggcccctga gcctcccacc caggtgggca ctttggaggc ccgagcaaca gccccacctg 420
ccccctcagc cccaaattct gctcccagca acttgggccc ccagaccgta ctggaggtcc 480
cagcccggag caccttctgg gggccccagc cctgggaggg gaggcccccc gccacaggcc 540
tggtgagctg ggctgaaccc gagcagaggc cagaggccag cgtccagttt gggagccccc 600
aggccaggag gcagcggcca gggagcccgg atcctgagtg gggcctccag ccacgggtca 660
ccttggagca gatctcagct ttctggaagc gtgaaggccg gaccagtgtg gggttctgaa 720
tccccagggt tccccagaga tccccgaggc aggccttgcc tcagtgggac cggggacccc 780
aggatccagc attaggattg agactgcccc agcgaagatg cccttcccag gctccttcca 840
cctggagtcc ccctccccgg gtctgggtgg tggccaggct atgtggacta ggggaagccc 900
agcagtgcct ctgctcagct acctgggctg tggctcagag acctgggggt ggagccaatg 960
ccaggccaga agccttcaag atcgcatcca gatgaagaac ccaaggtact agatagtcag 1020
gaaatggcat cgaccagcca cctccacctt ctttcagtgt ttaccgaagc caccaatacc 1080
aaagagaacg ggtcctgcgg tgctgaacag cctcggtgtg gcgatgacag ctggcaggag 1140
atgacaggaa tccagtttcc cagagccaca aatcctgttc tccttggcca ctcacccact 1200
gtgaggtcct ctaggaaaat acacaaagag aggaccagac caggcagagg aacattttgt 1260
ttcatatgaa ctgtggcttt gacccccaaa ctgcaaggag gaacttgctg ggccaagctg 1320
cagcggcact gtcttgctgg agtggggacc tagagtcaga gaaaacccac aggctcctct 1380
gcccattctc ctccatctgc acacgtctca gcctcggacc ctcaccactc catggtgagg 1440
aaggccatgg ccaggggaaa ctgagtttca tccaatgtgg agaggagcgt tgtcctagag 1500
cagggcaatc ccaaaatgtg acctctgatc atcgtccttt ccagcttg 1548